EFFECTS OF COMMON GENETIC VARIANTS IN *TP53* AND *TLR8* ON IMMUNE RESPONSE AND RISK OF CANCER IN PEOPLE WITH HIV/AIDS

by
Filip Pirsl

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science.

Baltimore, Maryland

August 2019

Abstract

Background: Cancer represents a significant source of morbidity and mortality in people living with HIV/AIDS (PLWHA) and risk of most cancer exceeds that observed in the general population. A substantial proportion of cancers in PLWHA may be attributable to infection, with risk largely due to HIV-related immune deficiency and co-infection with oncogenic pathogens. p53 is a tumor suppressor also involved in innate immune signaling via the Toll-like receptor 8 (TLR8). SNPs in both the *TP53* (rs1042522) and *TLR8* (rs3764880) genes may mediate the innate immune response, with the *TP53* G and *TLR8* A alleles hypothesized to be jointly protective against cancer.

Methods: Seven hundred and sixty participants enrolled in the Multicenter AIDS Cohort Study (MACS) provided blood samples for ascertainment of SNP genotypes. Outcomes assessed were diagnosis of any cancer, AIDS-defining cancer, non-AIDS-defining cancer, and infection-related cancer up to two years after participants' last visits. Exact Poisson regression was used to estimate unadjusted and adjusted incidence rate ratios associated with SNP genotypes for each outcome. Joint SNP effects were estimated using an interaction term.

Results: SNPs were found to be jointly protective against any cancer (interaction IRR: 0.46, 95% CI: 0.01-42.71), ADCs (interaction IRR: 0.47, 95% CI: 0.01-48.55), and infection-related cancers (interaction IRR: 0.40, 95% CI: 0.01-39.37) in multivariable models while main effects of SNPs were slightly protective or had no effect for all outcomes.

Discussion: These findings are consistent with a hypothesized synergistic effect of SNPs on the immune response. Weak main SNP effects and strong interaction indicate a protective effect of SNPs only in the presence of each other. Future work will address missing data using imputation and potential effects of sex by adding data collected by the Women's Interagency HIV Study.

Primary Reader: Eric C Seaberg, PhD

Secondary Reader: Christine Ladd-Acosta, PhD

Table of Contents

List of Tables

List of Figures

Introduction

HIV/AIDS and Treatments

Acquired Immune Deficiency Syndrome (AIDS) was first identified in the early 1980s in men who have sex with men afflicted with opportunistic infections and rare cancers usually associated with profound immunosuppression.[1] Not long after, multiple groups identified the causative agent to be the human immunodeficiency virus (HIV), a lentivirus of the retrovirus family.[2-6] As a member of the retrovirus family, HIV is an enveloped virus with an RNA genome that is reverse-transcribed to DNA by a viral reverse transcriptase enzyme.[6] The targets and mechanism of infection, namely T-cells expressing the CD4 surface antigen via either the CCR5 or CXCR4 co-receptor, were elucidated in the following years.[7-11] It was also established that HIV transmission occurs through mucosal contact with infected body fluids, most commonly via unprotected sexual intercourse as well as via maternal-infant transmission or transmission among injection drug users.[12,13] Upon infection, the virus is initially undetectable for up to several weeks, followed by rapid replication and increasing viral load, destruction and progressive decline of CD4+ cells, and resultant characteristic immunodeficiency.[12,13]

Since its emergence, HIV/AIDS has spread worldwide, infecting tens of millions and inflicting significant morbidity and mortality, with the highest burden present in sub-Saharan Africa.[13,14] The discoveries of HIV, its targets, and its mechanism of action spurred drug development, resulting in the identification and approval of several classes of drugs found to slow progression of disease.[15,16] The first drug approved for HIV in 1987 was azidothymidine (AZT), a nucleoside analog reverse-transcriptase inhibitor (NRTI), however, resistance was found to develop rapidly in those who responded to treatment.[17,18] In subsequent years, several additional drugs of the NRTI class were developed with no significant improvement in outcomes until the introduction of HIV-

protease inhibitors (PI), with saquinavir the first to be approved in 1995. The observation of resistance developing following treatment prompted investigations of combination treatments, with combination therapy of NRTI and PI found to be more effective than treatment with NRTI alone.[19] Not long after, the first non-nucleoside analog reverse transcriptase inhibitor (NNRTI), nevirapine, was approved and added to the growing list of combination therapies under investigation, culminating in the first regimen of highly active anti-retroviral therapy (HAART) in 1996.[20,21] The combination regimen constituting HAART has evolved over time as new more effective agents and classes, such as integrase inhibitors, have become available, but the initial introduction of HAART remains a monumental juncture in the HIV/AIDS epidemic.[22]

Global incidence peaked shortly after the introduction of HAART with 3.4 million new infections in 1996 and subsequently declined to 2.6 million new infections in 2005 and 1.8 million new infections in 2017.[13,14,23] Meanwhile, AIDS-related deaths peaked at 1.9 million deaths worldwide in 2004 and decreased to just under one million deaths in 2017.[13,23] Effective treatment for HIV has reduced incidence of new cases and mortality, resulting in a growing population of people living with HIV/AIDS (PLWHA) which numbered almost 37 million people in 2017.[23]

The introduction of effective combination therapy ushered in the HAART era and resulted in an epidemiologic shift in PLWHA, with significant reduction in deaths due to AIDS-related causes, prolonged survival, and with adequate access and adherence to HAART, the rise of the idea of HIV as a chronic disease.[13,24-27] However, with increasing survival and an aging population of PLWHA in the HAART era, it has been observed that this population is at elevated risk for numerous non-AIDS-related morbidities compared to the general population, including cardiovascular disease, cognitive decline, and cancer, among others.[28-32]

Cancer in People living with HIV/AIDS

Cancers represent a significant source of morbidity and mortality in PLWHA. Risks of most cancers exceed risk observed in the general population and cancer accounts for 20-30% of deaths in this population.[32-38] Elevated risk of cancer in PLWHA compared to the general population can, depending on the cancer type, be attributed to several factors, including frequent co-infection with oncogenic pathogens,[39-41] immune deficiency or dysregulation related to HIV,[42-44] and a higher prevalence of certain lifestyle-related risk factors for cancer in PLWHA, such as smoking,[38,45] alcohol use and abuse among certain sub-populations,[46,47] and potentially metabolic syndrome [48,49] which often arises due to HAART-related metabolic perturbations.[50,51]

Cancers in PLWHA are classified as AIDS-defining cancers (ADCs) or non-AIDS-defining cancers (NADCs) and demonstrate variable epidemiologic trends since the emergence of HIV and AIDS and their treatments.[52] The category of ADCs – consisting of Kaposi sarcoma (KS), non-Hodgkin lymphoma (NHL), and cervical cancer – is defined as such due to the early inclusion of these cancers in case definitions of AIDS.[1,53,54] The development of KS and cervical cancer in PLWHA occur as a consequence of coinfection with Kaposi sarcoma-associated herpesvirus (KSHV, also known as human herpesvirus-8) and human papilloma virus (HPV), respectively.[41,55] NHL, in contrast, is a more heterogeneous disease and different subtypes have been associated with several viral infections; including primary effusion lymphoma with KHSV,[56,57] multiple lymphoma subtypes with Epstein-Barr virus (EBV),[58,59] and marginal zone lymphoma and diffuse large B-cell with hepatitis C virus (HCV);[60,61] while many subtypes are often attributable to non-infectious factors that may involve lifestyle and occupational exposures.[62,63]

The category of NADCs consists of all cancers not included in ADCs and is not informative of etiology or pathogenesis of a given cancer. This is a heterogeneous category and includes both

cancers that have and have not been shown to have an infection-related etiology. Infection-related cancers among NADCs include cancers of the oral cavity, pharynx, and anogenital area excluding the cervix associated with HPV,[64,65] Hodgkin lymphoma associated with EBV,[66] liver cancer associated with viral hepatitis,[67] lung cancer and pneumonia,[68] and stomach cancer and *Helicobacter pylori*.[69] The remaining cancers in the category are non-infection-related cancers, some of which (e.g. breast, colorectal, ovarian, and prostate cancer) have been reported to occur less frequently among PLWHA.[70,71] Some of these findings remain topics of debate, such as reduced risk of prostate cancer being attributed to differential screening in separate studies,[72,73] while others, including the reduced risk of breast cancer, have taken root with potential mechanisms under investigation.[74]

Infection-related cancers straddle the demarcation of ADCs and NADCs, including cancers from both classifications, and may represent a more informative categorization than traditional ADC/NADC dichotomization. Given the setting of impaired immunity in this population, the development of these malignancies in PLWHA is an area of ongoing research and concern. The burden due to infection-related cancer in PLWHA is high, with as many as 40% of all cancers attributable to infection and almost 70% of cancers among PLWHA in their 20s attributable to infection.[75] Moreover, PLWHA are at significantly higher risk for infection-related ADCs and NADCs than the general population, although the degree of difference in risk has decreased in the HAART era.[76] The availability of measures to prevent or treat infections associated with these cancers makes this area an opportunity for public health intervention to reduce cancer burden in PLWHA.

In addition to its profound effects on disease progression and survival, HAART has also significantly influenced the epidemiology of cancer in PLWHA. The overall risk of cancer has

declined steadily since the introduction of HAART and cancer burden predominantly due to ADCs pre-HAART has given way to increasingly large burden due to NADCs in the HAART era.[37,77-80]

Soon after the effectiveness of HAART was first demonstrated, it was shown that clinical and survival benefits were mediated by suppression of HIV viral load and recovery of immune cells and function.[21,81] However, immune recovery may be incomplete among those initiating therapy with moderate or advanced disease[82-84] and long-term impairments, including chronic inflammation as well as immune deficiency and dysregulation, may persist despite successful viral suppression and CD4+ cell recovery.[85] The impaired immunological milieu in PLWHA may contribute to elevated risk of certain cancers, particularly infection-related malignancies, due to deficient immune response and reduced immune surveillance. A randomized trial comparing early initiation of HAART in asymptomatic HIV-infected individuals to deferment of initiation until a decrease of CD4+ cells below 350 cells per cubic millimeter found that early initiation resulted in a 60% reduction in hazard of any cancer, albeit at a median follow-up of only three years.[86] Secondary analysis of data from the same trial found a 74% reduction in hazard of infection-related cancer in early initiators compared to deferred initiators with suppression of viral load as a potential mediator of this relationship, again in spite of relatively short follow-up.[87] A similar trial with a slightly different definition for initiation of deferred treatment observed relatively few incident cancers and no reduction in risk among early initiators at a median follow-up of two years.[88] Observational studies have also shown that earlier initiation of therapy and better immunological response to therapy reduce risk of infection-related cancers.[42-44,89,90] Finally, a meta-analysis comparing PLWHA and solid organ transplant recipients found similar patterns of risks of cancers, suggesting that immune deficiency may confer risk in a similar manner in the two groups.[91]

In addition to immune deficits mediated by HIV and HAART, other modulators of immunity may contribute to risk of cancer among PLWHA. Non-heritable factors have been shown to be the primary determinants of immunity, however, genetic variants also contribute to immune function and can influence the immune response.[92,93] The human immune system can be separated into the non-specific innate immune system and highly-specific adaptive immune systems. The innate immune system represents the first line of defense in which pattern-recognition elements identify common motifs in pathogens and initiate an immune response via activation of the complement system, recruitment of immune cells, and activation of the adaptive immune system.[94] The adaptive immune system mounts a delayed but more robust and tailored response to pathogens through the generation of immune effector cells specific to antigens obtained from invading pathogens. This process also establishes immunological memory to enable a rapid adaptive immune response should infection with a given pathogen reoccur in the future.[94]

The human immune system is a complex network of regulatory mechanisms, signaling pathways, and cellular interactions, among a multitude of other elements. The sum of these parts is a dynamic and well-tuned framework able to respond to both internal and external threats.[94] Genetic variants in the immune system can disrupt or alter pathways within this network by affecting function or expression of its components, resulting in altered susceptibility to infection, autoimmunity, or immunodeficiency.[95,96]

p53

The transcription factor p53, most commonly known for its tumor suppressor functions and as a result dubbed the "guardian of the genome", has also been found to exert transcriptional regulation on components of the innate immune system, notably those involved in antiviral immunity.[97-101] Somatic genetic variants in the *TP53* gene are well-described in the setting of tumorigenesis, with

TP53 mutations identified in approximately half of all human cancers, while many germline variants have been found to be highly deleterious and associated with the Li-Fraumeni syndrome, a predisposition to early-onset cancer.[102,103] Non-Li-Fraumeni germline variants, which modify but do not remove function, may impact the role of p53 in transcriptional control and downstream regulation of innate immunity.

p53 was first discovered in *in vitro* studies of the oncogenic simian vacuolating virus 40 (SV40) and found to be the cellular target of SV40 large T antigen, the protein found to be necessary for SV40-mediated malignant transformation and subsequent tumor formation.[104,105] Following its identification, researchers cloned and investigated the role of p53 in tumorigenesis, with several experiments demonstrating the capacity of p53 to transform cells in combination with the previously identified *Ras* oncogene, prompting hypotheses of p53 as the protein product of an oncogene.[106-108] However, in subsequent years, it became clear that p53 proteins found to promote tumorigenesis were mutant forms and that the wildtype form is protective against tumors.[109-110] p53 was also found to be a cellular target of adenovirus E1 protein and the HPV E6 and E7 proteins with binding resulting in degradation of p53, establishing the role of p53 as a tumor suppressor by the end of the 1980s.[111-113] More recently, the LANA protein of KSHV has also been found to inhibit p53 function.[114] The targeting of p53 by viral proteins and subsequent degradation appears to constitute a viral mechanism enabling immune evasion, control of host cell machinery, and viral replication.[115,116] Conversely, intact p53 may be necessary for replication of some other viruses.[116]

p53 is a 53kD protein encoded by the *TP53* gene located on the long arm of chromosome 17 and functions as a sequence-specific DNA-binding transcription factor.[117,118] The *TP53* gene contains multiple sites subject to post-transcriptional alternative splicing resulting in as many as fifteen possible isoforms of the p53 protein with potentially variable functions.[119,120] The most abundant

isoform of p53 is a homotetramer of 393-amino acid peptides containing multiple domains which facilitate its broad interactions.[120,121] The regulatory function of p53 extends to numerous cellular processes, including cell cycle regulation, apoptosis, senescence, and immune response, among many others.[99-101,122,123]

The most extensively studied germline variant of *TP53* is the codon 72 single nucleotide polymorphism (SNP). In this SNP, officially designated rs1042522, a cytosine (C) base appears instead of a guanosine (G) base, resulting in a proline residue (P72) replacing an arginine residue (R72).[124] The location of this SNP is in a proline-rich region critical to p53's role in apoptosis and growth suppression and it has been shown to influence p53 interactions, with R72 having a greater affinity for nuclear export, trafficking to mitochondria, and induction of apoptosis.[125-129]

The mutant allele of rs1042522 is relatively common although allele frequencies appear to vary significantly across ancestral populations. Several studies have assessed the frequencies of the C/P72 and G/R72 alleles in different populations, finding that the frequency of the C allele increases incrementally with decreasing latitude, with prevalence of 15-25% among northern European populations, 35-55% in east Asians, 50-60% in south Asians, and 60-65% in equatorial African populations.[130-132] Given these findings, it has been suggested that some latitude or temperature-based selection has taken place, potentially related to the association of this SNP with fertility at different temperatures.[133,134]

Many population studies have assessed the effect of rs1042522 on risk of cancer with mixed results. Findings have included small suggestive association or no relationship with breast cancer[135-138] as well as associations of the C/P72 allele with lung cancer in Asian smokers, non-smokers, and a population of mixed smoking status;[139-141] C/P72 with lung cancer in African-American smokers,[142] G/R72 with lung cancer in a large consortium analysis,[143] C/P72 with

bladder cancer in Asians but not Caucasians,[144] C/P72 with endometrial cancer in Asian but not Caucasian women,[145] G/R72 with slightly reduced risk of gastric cancer among Asians,[146] and no association with risk of colorectal cancer.[147] In addition, the G/R72 allele has been found to have a greater susceptibility to degradation mediated by HPV E6 protein and may increase risk of oral but not penile carcinoma among those infected with HPV.[148-150] C/P72 has been associated with cervical cancer in east and south Asian populations, however, a pooled analysis of a number of studies found that risk was higher among carriers of the G/R72 allele.[151-153]

The collection of these findings does not present a very clear picture, although it is possible p53 plays variable roles in the tumorigenesis of different cancer types. Moreover, some have argued that such studies have been biased by limitations of study designs or ascertainment of genotypes, resulting in spurious associations or excess risk attributed to a specific allele.[153,154]

The growing appreciation for the breadth of the roles and interactions of p53 has prompted studies to determine its targets in the genome through the determination of and search for p53 response elements (REs). p53 REs are sequences that act as binding targets for p53 in the setting of transcriptional control and they can vary in affinity, number, and positioning with respect to a gene under p53 transcriptional control and its promoter region. In addition to characteristics of the p53 REs, numerous cofactors influence whether p53 will bind a given RE and if activation of transcription will take place.[155] The study of structural interactions in the binding of p53 to DNA are complex and various models have emerged, potentially reflecting the extensive diversity in pathways and interactions in which p53 engages. Many sequences have been validated and are considered canonical p53 REs while genome-wide analyses have suggested that many more potential binding sites exist.[156,157] The discovery of “half-sites” of ten bases and “three-quarter

sites" serving as p53 binding sites in a similar manner to low-affinity canonical 20-base RE sequences has also significantly expanded the number of potential targets.[158,159]

The ability and affinity of p53 binding to its REs can be altered by mutations in p53 resulting in various downstream effects.[160] The same can also be said for mutations in the RE sequences themselves with numerous SNPs shown to disrupt functional p53 REs. Among these, a mutation in the promoter region of Toll-like receptor (TLR) 8 of the innate immune system has emerged as a SNP of interest.[161,162]

p53, Toll-like Receptor 8, and the Innate Immune Response

The Toll-like receptors (TLRs) are a family of microbial sensors involved in the innate immune system's non-specific identification of pathogens via recognition of pathogen-associated molecular patterns. TLRs are transmembrane proteins located either at the cell surface or in intracellular endosomes and contain a sensory extracellular or luminal domain and an intracellular signaling domain. The members of the TLR family act as sensors for different microbial components and types of pathogens, including both bacteria and viruses, and initiate different pathways and responses upon stimulation.[94,163] Additional roles of TLRs have been identified in other normal cellular processes, including cell differentiation and survival, as well as disease, including autoimmune disorders and cancer.[164-166]

The TLR-dependent innate immune response has been described as having been thought of as "hard-wired" with gene induction almost exclusively mediated by interactions of TLRs with their respective stimuli, however, more recent studies have challenged this view with findings indicating that p53 influences expression of TLRs.[162,167-170]

TLR8 is an intracellular sensor of single-stranded viral RNA and induces production of antiviral interferons upon stimulation.[94,171] The SNP rs3761624, featuring either an adenosine or guanosine base, is located in the promoter of *TLR8* within a p53 RE and determines the binding affinity of p53 for this site.[161,162] rs3761624 has been observed to be in perfect linkage disequilibrium with rs3764880, a nearby SNP in exon 1 of *TLR8* characterized by a guanosine base replacing an adenosine base, which is more readily available in SNP panels and has received more attention in the literature.[172] The G-allele of rs3764880 has previously been associated with slower progression of HIV in studies of mostly white and mostly east African participants but also with higher peak HIV viral load in a study of a Kenyan perinatal HIV cohort[173-175] while the A-allele has been associated with reduced susceptibility to tuberculosis.[176-178] TLR8 is encoded by a gene located on the X chromosome, meaning effects may be modified by sex, such as via X-inactivation in women.[177] According to 1000Genomes, frequencies of the A allele of rs3764880 are 73% in white European, 71% in African, 19% in east Asian, and 42% in south Asian ancestral populations with nearly identical frequencies for the A allele of rs3761624.[179,180]

Germline genetic variants in *TP53* and *TLR8* may individually or jointly influence risk of oncogenic infection and cancer in PLWHA through their effects on the p53-immune response signaling axis. It has been established that PWLHA are at elevated risk for infection-related cancers compared to the general population and additional non-HIV deficits in the immune response may confer additional risk.[76]

The effect of *TP53* SNP rs1042522 on its ability and affinity to bind to its REs has received little attention. One study assessed the effect of this SNP on p53 interaction with NF-κB, another major transcription factor, and transactivation of genes involved in immunity inflammation and found that the P72 allele promoted stronger interactions with NF-κB and a stronger immune response to

stimulation with lipopolysaccharide, a component of many bacterial cell walls.[181] The findings from this study suggest that rs1042522 likely does influence the affinity of p53 for some REs. It is not known how rs1042522 influences p53 binding with the RE in the promoter of *TLR8* in the presence or absence of rs3764880, but it is likely that there is modification of the binding affinity by both SNPs.

The primary objective of this study is to assess the main and joint effects of rs1042522 and rs3764880 on incidence of all cancers, ADCs, NADCs, and infection-related cancers. Based on prior findings in studies of these SNPs as well as preliminary analysis of their effects on risk of AIDS following HIV seroconversion, we hypothesize that the G/R72 allele of *TP53* and A-allele of *TLR8* act synergistically to reduce risk of cancer in a cohort of HIV-positive men. The findings of this study could potentially aid in stratifying PLWHA by risk of cancer and guiding screening strategies accordingly.

Methods

Study Population

This study was conducted in a subset of participants of the Multicenter AIDS Cohort Study (MACS). The MACS is an ongoing prospective cohort study of HIV-infected and HIV-uninfected men which began with an initial recruitment of participants in Baltimore, Maryland; Chicago, Illinois; Los Angeles, California; and Pittsburgh, Pennsylvania from April 1984 through March 1985. Subsequent waves of enrollment emphasized recruitment of minority participants and took place from April 1987 through September 1991 and October 2001 through August 2003. More recently, ongoing enrollment since 2010 has recruited participants to study effects of modern

therapies.[182-184] The MACS was approved by the Institutional Review Boards at all study sites and all participants provided informed consent.

Participants attended study visits by protocol approximately every six months and underwent physical examination, blood sample collection, neuropsychological evaluation, and questionnaires assessing medical history, behavior, quality of life, depression, activities of daily living, and medication use at each visit. Aliquots of blood samples were used to assay biomarkers of disease, including CD4+ cell count and HIV viral load, and other biological measures of immediate interest.[184] Blood samples were also processed and aliquoted and specimens were kept locally for site-specific assays or sent to a central repository to be made available to investigators for future research. AIDS and non-AIDS diagnoses were ascertained by participant self-report and subsequent confirmatory medical record review or from other sources, such as death certificates as well as cancer and tuberculosis registries.[183,185]

This study included participants enrolled in the first three waves with available SNP data and who contributed person-time under observation while HIV-positive. These participants were selected to explore the association of host genetics on HIV susceptibility and progression.

Outcomes of Interest

The primary outcome of interest in this study was incident cancer. Cancer diagnoses and dates of diagnosis were obtained by participant self-report at study visits with subsequent confirmation via medical record review or captured using state cancer registry data or death certificates.[182] Cancer site and histology were documented using National Cancer Institute Surveillance, Epidemiology, and End Results Program (SEER) cancer site recodes for the International Classification for Diseases for Oncology, third edition (ICD-O-3).[186] Non-melanoma skin cancers, which are not

documented by state cancer registries and are not coded by SEER, were not considered in this analysis.

Multiple classifications of cancer were considered as outcomes in this study, namely any cancer, ADCs, NADCs, and infection-related cancers. The outcome of any cancer was defined as the first occurrence of diagnosis of cancer in any SEER-defined site. ADCs were defined as the first occurrence of diagnosis of NHL, either nodal (SEER cancer site code: 33041) or extranodal (33042), or KS (36020). NADCs were defined as the first occurrence of diagnosis of cancer in a SEER-defined site not included in the definition of ADC, specifically any SEER code other than codes 33041, 33042, and 36020. Lastly, infection-related cancers were defined as the first occurrence of diagnosis of cancer in any of the following sites: oral cavity and pharynx, including lip (20010), tongue (20020), salivary gland (20030), floor of mouth (20040), gum and other mouth (20050), nasopharynx (20060), tonsil (20070), oropharynx (20080), hypopharynx (20090), and other oral cavity and pharynx (20100); stomach (21020); anus, anal canal and anorectum (21060); liver (21071), larynx (22020), lung and bronchus (22030); Hodgkin lymphoma, either nodal (33011) or extranodal (33012); NHL, either nodal (33041) or extranodal (33042); or KS (36020). Participants were able to contribute events to multiple cancer classifications and were not censored at time of diagnosis of a cancer that was not included in a particular outcome definition. As a result, a cancer diagnosis could contribute an event to one analysis but not another, even if it satisfied the outcome definition, if it was a participant's second cancer diagnosis. Participants who experienced multiple events of the same cancer classification only contributed the first diagnosis as an event to analysis.

Events were included in the analysis if they occurred while participants were under or within two years following a participant's most recent study visit. Events that occurred more than two years

following a participant's most recent preceding study visit were excluded from consideration and participants were censored at the date of the most recent study visit preceding the diagnosis. Participants who did not experience an event were censored at the date of their most recent study visit. Remaining ongoing participant follow-up was administratively censored on December 31, 2013 to allow for delays in ascertainment of new cancer cases.

Exposures, Covariates of Interest

The exposures of interest in this study were the SNPs rs1042522 in *TP53* and rs3764880 in *TLR8*. Specimens stored in the MACS central repository were used to run SNP panels to provide investigators with genetic data for the study of effects of polymorphisms on a variety of outcomes. Genotypes for SNPs rs1042522 and rs3764880 were obtained from data generated by this effort in three separate initiatives, which included the use of TaqMan and Illumina-based genotyping assays and standard quality control measures. rs3764880 has been reported to be in perfect linkage disequilibrium with rs3761624,[172] the SNP located in the p53 RE of the *TLR8* promoter, and was used as a surrogate for the latter. rs1042522 genotypes, assessed in one genetic testing initiative, were determined only for a subset of participants who were evaluated for rs3764880, which was assessed in three genetic testing initiatives in the MACS. The possible genotypes at rs1042522 were CC, CG, and GG. Since participants of the MACS are exclusively men and the gene encoding TLR8 is located on the X chromosome, participants only had one copy of *TLR8* and one allele at rs3764880, thus, the possible genotypes at rs3764880 were G* and A*. SNP genotype data were linked to data collected at study visits as well as data on cancer diagnoses and were assessed for individual main effects of the *TP53* SNP rs1042522 (GG vs. CG/CC genotype) and the *TLR8* SNP rs3764880 (G* vs. A* genotypes) as well as for a hypothesized joint protective effect of the

rs1042522 GG and rs3764880 A* genotypes. The joint effect was modeled using an interaction term of the two hypothesized protective genotypes.

Covariates of interest were collected at study visits and consisted of demographic data, including age and self-reported race; lifestyle and behavioral characteristics, including body mass index (BMI; kg/m^2), current smoking status (never, current, or former), pack-year smoking history since last visit, cumulative lifetime pack-year smoking history, and alcohol consumption in terms of number of drinks consumed per week; and measures of HIV and its treatments, including CD4+ cell count (cells/mm^3), plasma HIV viral load (copies/mm^3), and current antiretroviral therapy (none, monotherapy, non-HAART combination therapy, or HAART). Antiretroviral therapy was defined as HAART based on DHHS/Kaiser Panel guidelines as three or more antiretroviral drugs consisting of one or more PIs, or one NNRTI, or one nucleoside or nucleotide reverse transcriptase inhibitor, or an integrase inhibitor, or an entry inhibitor (including fusion inhibitors).[184] Finally, treatment eras were also considered, including the pre-HAART (1984 through 1995) vs. HAART (1996 through end of follow-up) eras as well as the pre-HAART vs. early HAART (1996-2001) vs. modern HAART (2002 through end of follow-up) periods.

TP53 SNP rs1042522, *TLR8* SNP rs3764880, and self-reported race were treated as fixed variables while age at visit, BMI, smoking status, pack-year smoking history since last visit, cumulative lifetime pack-year smoking history, alcohol consumption, injection drug use status, CD4+ cell count, HIV viral load, and antiretroviral therapy at time of visit were treated as time-varying variables. Body mass index, smoking status, pack-year smoking history since last visit, cumulative lifetime pack-year smoking history, alcohol consumption, injection drug use status, CD4+ cell count, HIV viral load, and antiretroviral therapy at time of visit were lagged two years to capture

more etiologically relevant measures of these variables and reduce the potential for reverse causality in covariate-exposure and covariate-outcome relationships.

Missing data for exposures and covariates of interest were addressed using multiple methods. Missing data for time-varying covariates were imputed using a last observation carried forward approach. Observations were carried forward for up to four visits, corresponding to two years. The assumption of this approach was that true values for these covariates were stable and adequately captured by the last non-missing observation carried forward. Missing values for antiretroviral therapy remaining in the first ten scheduled MACS study visits corresponding to the first five years of the study (1984-1989) were imputed as participants not receiving antiretroviral therapy given that the first effective treatment was not approved until 1987 and participants with missing data in the two years following this date were assumed to have not yet initiated treatment. Similarly, remaining missing values of HAART for study visits preceding 1996 were imputed as not receiving HAART since such regimens were not available until this year. Missing values of antiretroviral therapy and HAART at study visits that preceded participants' HIV diagnoses were imputed as not receiving therapy or HAART with the assumption that they were not receiving prophylactic antiretroviral therapy.

For the purposes of analysis, covariates were categorized as follows: age at visit (less than 35, 35-45, 45-55, or over 55), race (white, black, Hispanic, or other), CD4+ cell count (less than 200, 200-350, 350-500, or over 500 cells/mm^3), HIV viral load (less than 1000 or 1000 or greater copies/mm^3), cumulative pack-year smoking history (less than 10, 10-20, 20-30, or over 30), alcohol consumption in terms of number of drinks consumed per week (none, 1-3, 4-13, or over 13), and BMI (less than 25, 25-30, or greater than 30). In addition, the HAART and HIV viral load covariates were combined with the following categories: not receiving HAART, receiving

HAART and but not virally suppressed (HIV viral load of 1000 copies/mm^3 or greater), and receiving HAART and virally suppressed (HIV viral load under 1000 copies/mm^3). During analysis, covariates were collapsed into fewer groups if categories showed similar associations with outcomes.

Statistical Analysis

Demographic and baseline characteristics were described as ascertained at participants' baseline visit. Participants contributed person-time to this analysis only while HIV-infected, meaning those who were HIV-infected at enrollment began contributing person-time at this visit while those who were HIV-uninfected at enrollment did not contribute person-time until seroconversion. Person-time contribution ended at the time of an incident event, time of censoring, or death. Person-time contribution did not end if a participant was diagnosed with a cancer not included in the outcome definition.

Cancer incidence rates (IR) were calculated by dividing the number of events observed by the person-time of follow-up accrued and exact Poisson regression was used to estimate incidence rate ratios (IRR). The exact method was used to compute confidence intervals due to small numbers of events in combinations of strata of exposures and covariates of interest which make maximum-likelihood estimate-based Poisson regression unsuitable.

Univariable exact Poisson regression was used to estimate crude IRRs for the SNPs and covariates of interest. Covariates were then individually modeled with main and joint SNP effects to assess influence of adjusting for each covariate on SNP estimates for consideration during multivariable modeling. Multivariable exact Poisson regression was conducted to estimate adjusted IRRs for SNP effects. Self-reported race, used to adjust for ancestry as well as effects attributable to race,

was forced into the multivariable model for each outcome. Additional covariates included in the final multivariable model were determined based on available knowledge of associations with outcomes and on the extent to which covariates altered SNP effects when modeled individually with SNPs. All analyses were performed in Stata 15.[185]

Results

Study Population

Our study population included 760 MACS participants with available SNP data and study visit data. At enrollment, 318 participants (41.8%) were HIV-infected and 442 participants (58.2%) were HIV-uninfected. Among participants who were HIV-uninfected at enrollment, the median time from enrollment to seroconversion was 2.5 years (IQR: 1.0-6.1) and median age at seroconversion was 35.1 (IQR: 29.6-41.6). Participants were predominantly enrolled in the first wave of MACS enrollment, with 626 participants (82.4%) enrolled in this period. Participants contributed a median of 9.3 years (IQR: 4.7-21.3) years of follow-up while HIV-infected. Ten participants contributed no follow-up while HIV-positive and were not included in time-to-event analysis.

All 760 participants had data available for *TLR8* SNP rs3764880, 222 (29.2%) of whom had the G allele and 538 (70.8%) of whom had the A allele at this site. 263 participants (34.6%) had data available for *TP53* SNP rs1042522, 115 (43.7%) of whom had genotype CC, 110 (41.8%) with genotype CG, and 38 (14.4%) with genotype GG at this site. Among participants with data for the *TP53* SNP, 213 participants (81.0%) were enrolled in the first wave of MACS enrollment. In

addition, all participants in this subset were HIV-positive at enrollment and contributed a median 8.9 years of follow-up (IRQ: 4.9-25.5).

Baseline characteristics

Baseline characteristics obtained at enrollment visit were relatively evenly distributed when participants were stratified by *TLR8* SNP allele (**Table 1**). Differences were observed in smoking status, with more current smokers among participants with the G allele, and HIV status, with more HIV-positive participants among participants with the A allele.

However, stratifying by *TP53* SNP genotype revealed several differences in baseline characteristics between the groups (**Table 2**). The distributions of self-reported race varied across genotypes, with clear trends in proportions of participants reporting white, black, and Latin/Hispanic ancestry in each group. Moreover, differences in distributions were observed across *TP53* SNP genotypes in most covariates considered, some of which may be related to the differential allele and genotype frequencies observed by race. These included smoking history, cumulative pack-years of smoking history, and alcohol consumption.

The distributions of self-reported race by *TLR8* SNP status were similar and allele frequencies were relatively consistent across white, black and Latin/Hispanic participants. In contrast, allele frequencies of the *TP53* SNP varied by race, with the lowest frequencies of the protective G allele participants of Latin/Hispanic, other, and white groups while those with black ancestry had the highest frequency of this allele (**Table 3**).

Cancer Diagnoses

One hundred thirty-three first diagnoses of any cancer were observed during 8,734.0 years of follow-up, resulting in an incidence rate of 1.52 (95% CI: 1.28-1.80) first cancer diagnoses per

100 person-years. Cancer diagnoses occurred a median of 6.0 years (IQR: 4.6-10.2) after enrollment. The most common diagnoses were KS, NHL, and prostate cancer with ADCs (KS and NHL) accounting for two thirds of first cancer diagnoses (**Table 3**). The cumulative incidence curve of first diagnosis of any cancer shows greatest risk in the period of four to eight years after enrollment, followed by a sustained risk for the rest of follow-up (**Figure 1**).

In analysis of ADCs alone, 92 first diagnoses were observed during 8,905.9 years of follow-up time, resulting in an incidence rate of 1.02 (95% CI: 0.83-1.25) first ADC diagnoses per 100 person-years. Diagnoses of ADC occurred a median of 6.34 years (IQR: 4.2-7.6) years after enrollment. KS accounted for 80% of diagnoses, followed by extranodal and nodal NHL, respectively (**Table 3**). The cumulative incidence curve of first diagnosis of ADC shows greatest risk between four and eight years after enrollment with, with risk dissipating by twelve years and plateauing for the remainder of follow-up time (**Figure 2**).

In analysis of NADCs alone, 49 first diagnoses were observed during 8,949.7 years of follow-up time, resulting in an incidence rate of 0.50 (95% CI: 0.38-0.67) first NADC diagnoses per 100 person-years. Diagnoses of NADC occurred a median of 13.7 years (IQR: 6.0-18.3) after enrollment. The most common diagnoses were prostate cancer, cancers of the anus, anal canal and anorectum, and melanoma, respectively (**Table 3**). The cumulative incidence curve of first diagnosis of NADC suggests a relatively constant risk over the duration of follow-up time (**Figure 3**).

Finally, in analysis of infection-related cancers, 110 first diagnoses were observed during 8,847.6 years of follow-up, resulting in an incidence rate of 1.20 (95% CI: 0.99-1.45) first diagnoses of infection-related cancer per 100 person-years. Diagnoses of infection-related cancers occurred a median of 5.7 years (IQR: 4.5-8.2) after enrollment. ADCs accounted for most infection-related

cancers, followed by cancers of the anus, anal canal and anorectum (**Table 3**). The cumulative incidence curve of first infection-related cancer shows a similar trend to the curve for ADCs in the first 12 years of follow-up but does not plateau as infection-related NADCs continue to occur later in follow-up (**Figure 4**).

Univariable Analysis

The incidence of cancer of any type decreased over time, with a substantial reduction in risk in the HAART era observed compared to the pre-HAART era. Within the HAART era, risk of cancer overall appeared to remain relatively consistent, with no further reduction in risk in the modern compared to the early HAART era. Protective effects were observed for both SNPs, with 41% (IRR: 0.59, 95% CI: 0.15-1.63) and 17% (IRR: 0.83, 95% CI: 0.57-1.23) reduction in incidence of any cancer in participants with the protective genotypes of the *TP53* and *TLR8* SNPs, respectively (**Table 4**). Among covariates of interest, age did not show consistent association with cancer while non-white participants were found to have lower cancer incidence than white participants. Among measures of disease severity and therapy, higher CD4+ cell count and treatment with HAART were found to be protective against cancer diagnosis of any type. Risk was higher among participants with HIV viral load of 1000 copies/mm^3 or greater compared to those with lower HIV viral load, but this difference dissipated among those receiving HAART. Lifestyle factors; including smoking status, cumulative pack-year smoking, and alcohol consumption; showed inconsistent associations with cancer across categories.

Significant trends over calendar time were also observed for ADC, with an even greater reduction in risk observed for ADCs in the HAART era compared to the pre-HAART era than was observed for all cancers. Within the HAART era, incidence of ADCs was further reduced in the modern compared to the early HAART era. Protective effects were also observed for both SNPs with

respect to incidence of ADCs, with a 41% (IRR: 0.59, 95% CI: 0.11-1.91) reduction in incidence for participants with the protective *TP53* genotype and a 17% (IRR: 0.83, 95% CI: 0.53-1.33) reduction in incidence in participants with the protective *TLR8* genotype. Incidence of ADCs was found to decrease with increasing age, potentially due to lower engagement in high-risk behaviors among older participants, and risk was lower among black and Latin/Hispanic participants compared to white participants. Higher CD4+ cell count was found to be protective while higher HIV viral load was associated with ADC. No events were observed among participants receiving HAART, regardless of viral load. Participants consuming any alcohol and those who were underweight or normal per BMI were found to be at increased risk of ADCs, although confidence intervals included the null, while smoking status and cumulative pack-year history showed inconsistent trends, respectively (**Table 5**).

In contrast with all cancers and ADC, the incidence of NADC increased, with greater than two-fold higher incidence in the HAART than the pre-HAART era. Both SNPs were found to be protective, with 54% (IRR: 0.46, 95% CI: 0.01-3.13) and 21% (IRR: 0.79, 95% CI: 0.42-1.54) reductions in participants with the protective *TP53* and *TLR8* SNPs, respectively. Furthermore, unlike the results for all cancers and ADCs, older age was strongly associated with risk of NADCs. Black participants were at slightly lower risk for NADCs than white participants while no NADCs were observed for participants of Latin/Hispanic or other race. Higher CD4+ cell count was protective against NADC, however, confidence intervals for most estimates included the null, and HIV viral load showed no association with risk of NADCs. Incidence of NADCs was higher among participants treated with HAART, most likely due to prolonged survival allowing sufficient time for the development of cancers with longer latency. Interestingly, BMI showed reduced risk among overweight but increased risk of NADC among obese participants compared to underweight or

normal participants, however, confidence intervals for both estimates included the null. While smoking status showed inconsistent associations across categories, higher cumulative pack-years of smoking was associated with increased risk of NADC. Surprisingly, alcohol consumption did not show any association with NADC (**Table 6**).

Lastly, infection-related cancers also decreased over time, with a reduction in risk from pre-HAART, through early HAART, and into the modern HAART era. The effects of SNPs were again found to be protective, with 53% (IRR: 0.47, 95% CI: 0.09-1.48) and 13% (IRR: 0.87, 95% CI: 0.57-1.35) reductions in incidence in participants carrying the protective *TP53* and *TLR8* SNP genotypes, respectively, compared to non-carriers. Older age was associated with an increasingly protective effect, although this may again potentially be due to reduced high-risk behaviors in these participants. Incidence of infection-related cancers was lower among non-white participants compared to white participants. Increasing categories of CD4+ cell count showed an increasingly protective effect. Among participants receiving HAART, those with high viral load were at higher risk than those with low viral load, but at substantially lower risk than those not receiving HAART. Among lifestyle factors, incidence of infection-related cancer was lower among overweight or obese participants by BMI and higher in participants consuming any alcohol, although confidence intervals included the null. Smoking status and cumulative pack-years smoking were not associated with risk of infection-related cancer (**Table 7**).

The SNPs of interest both showed expected effects, with the GG genotype of *TP53* SNP rs1042522 and, to a lesser extent, A* genotype of *TLR8* SNP rs3764880 both found to be protective for all outcomes, although confidence intervals included the null for all three SNP estimates for each outcome.

Multivariable Analysis

For the multivariable analysis, we began by fitting a model to estimate the main and joint effects of the *TP53* and *TLR8* SNPs on the incidence of any cancer adjusted for race, BMI, smoking history, CD+4 cell count, and a composite measure of HIV treatment (HAART) and treatment effectiveness (HIV viral load). The main effects of both SNPs remained protective against cancer diagnosis of any type, with reduction in incidence of 35% (IRR: 0.65, 95% CI: 0.01-5.74) and 8% (IRR: 0.92, 95% CI: 0.33-2.93) for carriers of the protective *TP53* GG and *TLR8* A* genotypes, respectively, compared to non-carriers. These were slightly attenuated compared to effects observed in univariable analysis. The interaction between the two SNPs was found to be additionally protective, indicating a joint effect of the two SNPs. Participants with protective genotypes at both SNPs appeared to be at lower risk of cancer with the interaction of the two SNPs indicating a half-fold reduction in risk (IRR: 0.46, 95% CI: 0.01-42.71). Confidence intervals for all SNP effects, however, contained the null (**Table 8**).

Among covariates included, CD4+ cell count above 350 cells/mm^3 and HAART irrespective of viral load were found to protect against cancer. Smoking and overweight/obese BMI were also consistent with a protective effect, although these may be reflections of disease severity or overall health rather than true effects of these exposures on the outcome. Confidence intervals for effects of all covariates contained the null value except for that of CD4+ cell count (**Table 8**).

Multivariable analysis produced a similar model estimating main and joint effects of the *TP53* and *TLR8* SNPs on incidence of ADC with the same adjustment set of race, BMI, smoking history, CD+4 cell count, and the composite of HAART status and HIV viral load. The protective effect observed for the *TP53* SNP GG genotype in univariable analysis was entirely adjusted away, resulting in a null effect in the multivariable model, while the protective effect observed for the

TLR8 SNP A* genotype remained virtually unchanged, with a 16% reduction in incidence of ADC in carriers compared to non-carriers. The interaction between the two SNPs was found to be additionally protective, again indicating a joint effect of the SNPs on incidence of ADCs. Participants with protective genotypes at both SNPs were at lower risk for ADCs, with a 53% reduction (IRR: 0.47, 95% CI: 0.01-48.55) in incidence compared to participants not carrying both protective genotypes. The null association of *TP53* and weak protective effect of *TLR8* with a strong interaction indicate a protective effect of each SNP only in the presence of the other SNP. Confidence intervals for SNPs again all contained the null as was the case for the outcome of any cancer (**Table 9**).

Among covariates included, CD4+ cell count above 350 cells/mm^3 was protective against ADCs, albeit to a lesser extent than the effect observed for any cancer, but HAART alone was not. Among participants receiving HAART, those with viral load above 1000 copies/mm^3 were at slightly higher risk of ADC than those not on HAART while those with viral load below this threshold were at substantially lower risk. Smoking and overweight/obese BMI were found to protect against ADCs to a greater extent than observed for any cancer, again potentially reflecting disease severity or overall health of participants. Confidence intervals of covariate effects all contained the null value (**Table 9**).

Multivariable analysis was unable to produce a model estimating adjusted main and joint effects of SNPs on incidence of NADCs. Only one NADC was observed among participants with the protective genotypes for both SNPs, resulting in adjusted models including the interaction term failing to converge. Consequently, no adjusted SNP effects are reported for NADCs.

Finally, multivariable analysis produced a model estimating main and joint effects of SNPs on incidence of infection-related cancers adjusted for race, BMI, smoking history, alcohol

consumption, CD4+ cell count, and the composite of HAART status and HIV viral load. The protective effects for both SNPs observed in univariable analysis were attenuated following adjustment, resulting in a slightly protective effect for the *TP53* SNP (IRR: 0.86, 95% CI: 0.02-8.00) and a null effect for the *TLR8* SNP (IRR: 1.01, 95% CI: 0.33-3.69) in the final multivariable model. Incidence of infection-related cancer was reduced by 14% in participants with the *TP53* GG genotype compared to those carrying other genotypes while there was no difference in incidence based on *TLR8* SNP status. Participants with protective genotypes at both SNPs were found to be at lower risk, with a 60% (IRR: 0.40, 95% CI: 0.01-39.37) reduction in incidence compared to participants not carrying both protective genotypes. Similar to ADCs, the observation of weak main effects and stronger joint effect indicate SNP effects only in the presence of each other. Confidence intervals for SNP effects all contained the null (**Table 10**).

Similar to associations observed in the analysis of ADCs, CD4+ cell count over 350 cells/mm^3 was found to be protective while HAART alone was not. A protective effect of HAART was only observed in participants with viral load below 1,000 copies/mm^3 while participants with viral load above this threshold were at slightly increased risk compared to those not receiving HAART. Overweight/obese BMI and smoking were also found to be protective against infection-related cancers, again potentially reflecting disease severity or overall health, while greater alcohol consumption was found to increase risk. Confidence intervals for all covariate effects, with the exception of CD4+ cell count, included the null (**Table 10**).

Discussion

Significant attention has previously been devoted to the effect of the *TP53* SNP rs1042522 on risk of cancer,[135-154] however, to our knowledge this is the first study to assess its effect in combination with *TLR8* with respect to cancer. While p53 is best known for its tumor suppressor function, an increasing appreciation for the role of p53 in immune function has emerged, which may include a pathway through a p53 response element in *TLR8*, with binding subject to the *TLR8* SNP assessed in this study.[161-162] Accordingly, the significance of studying these SNPs as a pair is the potential for differential immune response by SNP status and a synergistic effect of the protective genotypes, which may identify participants at higher risk for development of cancer or may improve understanding of the etiology of certain cancers in PLWHA.

The results of this study are consistent with the hypothesis of a synergistic protective effect of SNPs in *TP53* and *TLR8* on the incidence of multiple classifications of cancer in MACS participants. The interaction between SNPs was found to be protective in multivariable models for any cancer, ADCs, and infection-related cancers – with strongest effect observed for the latter – while main effects of SNPs varied by outcome, with the *TP53* GG genotype shown to be most protective for incidence of any cancer and to have a null effect for ADCs while the *TLR8* A* genotype was most protective against ADCs and showed a null effect for infection-related cancers. Statistical interactions, representing the joint effect of SNPs, were observed to be protective in multivariable models for any cancer, ADCs, and infection-related cancers. These results support the hypothesized role of the p53-TLR8 axis in the immune response on the development of these cancers in men with HIV/AIDS.

Allele frequencies of *TP53* variants based on self-reported race differed in this study from those reported in the literature.[130-132] Studies have previously reported a lower frequency of the C/P72

allele among white northern European samples compared to samples taken from African populations and an inverse relationship between C/P72 allele frequency and latitude in east Asia, however, this study observed a higher frequency of the C/P72 allele among white and Latin/Hispanic participants than black participants.[130-132] There are several potential explanations for these observations, including relatively small sample size, discrepancy between genetic ancestry and self-reported race in the study sample, or possibly some sort of selection. Future avenues of research could address some of these potential explanations. Allele frequencies observed for *TLR8* SNP rs3764880 were concordant with those reported by 1000Genomes.[179,180]

SNP effects observed were similar across outcomes of any cancer, ADCs, and infection-related cancers, most likely due to the composition of cancers within these classifications observed during follow-up. ADCs were the predominant type of cancer observed in this study, accounting for 67% of events in the any cancer outcome and 85% of infection-related cancers. As a result, associations between SNPs and ADCs likely contributed significantly to effects observed for the other outcome classifications which included ADCs and may explain why final adjusted models were similar. Although this study did not have enough events to assess additional cancer classifications, it may be of interest to conduct analyses excluding ADCs to characterize the effects of these SNPs on general cancer development in both HIV and non-HIV populations.

Unfortunately, adjusted main and joint SNP effects on incidence of NADC could not be estimated in multivariable analysis. Only one NADC was observed among participants with protective genotypes at both SNPs and multivariable exact Poisson regression modeling failed when the SNP interaction term was included. While the unadjusted SNP effects for NADCs were similar to those observed for other cancers, the unadjusted effects of some other covariates differed. In the event that a multivariable model could be estimated, these differences could result in different estimates

of SNP effects from those observed for other cancer outcomes due to potential differences in how inclusion of covariates in a given model could influence estimated SNP effects.

Effects observed for other participant characteristics were largely consistent with previously reported associations,[38,42-49,86-90] with the exception of smoking. Smoking lagged by two years; categorized as never, ≤30 pack-year history, and >30 pack-year history; was associated with lower cancer risk in all multivariable models, with greater protection among participants with more extensive exposure. In this case, smoking may represent better overall health or less severe HIV disease, although this is not clear. Smoking showed a null or adverse effect on risk in univariable model and given previously identified effects of rs1042522 on apoptosis, which can occur as a consequence of cigarette smoke-induced damage,[127,128,187,188] it is possible that this SNP is also associated with smoking or its effects which was not fully accounted for in our modeling.

Similarly, participants who were overweight or obese by BMI were found to be at reduced risk compared to normal or underweight participants. This association may be explained by low BMI acting as a marker of disease severity through disease-related wasting, representing a high-risk group for cancer. Direct measures of disease severity and treatment, CD4+ cell count and a composite of HAART and viral load, both described associations of adequately treated and less severe disease with reduced risk of cancer, reiterating the importance of immune function in the development of cancer among PLWHA.

There were a number of limitations in this study which will be addressed in further work and analyses using these data, most notably using an all-male sample, missing data, and potentially inadequate adjustment for genetic ancestry. This analysis was conducted on data obtained from participants of the MACS, a study which exclusively enrolled male participants. Given differential risks of cancer by sex in the general population, it is likely that differences in risk by sex also exist

among PLWHA and possible that the effects of the SNPs of interest may be influenced or modified by sex – something that could not be assessed in this study population. This study also could not assess several cancers only observed in women, most notably cervical cancer, which is an ADC.

To address this limitation, further work on this analysis will incorporate data from the Women's Interagency HIV Study (WIHS). The WIHS is an ongoing prospective cohort study enrolling HIV-positive women which began with an initial recruitment of participants at six sites from October 1994 to November 1995 with multiple subsequent waves of enrollment at variable numbers of sites.[189] The conduct of and data collected by the WIHS study are similar to the MACS in terms of exposures, covariates, and outcomes of interest in this study, allowing for a combined analysis. There are, however, notable differences between the participants of the two studies. Some important differences are racial composition; with a greater proportion of black participants in the WIHS; and method of HIV transmission, with more exposures through IDU than sexual transmission in the WIHS than the MACS. These differences will likely require additional considerations in the analysis.

This study was also limited by missing data in exposures and covariates of interest, which substantially reduced analytic sample size. Out of 760 participants included in analysis, all participants were assed for the *TLR8* SNP in three separate genetic testing initiatives while 497 (65%) were not tested for the *TP53* SNP in any of these three studies, resulting in their exclusion from univariable and multivariable models including terms for this SNP.

Baseline characteristics among those missing *TP53* data varied compared to those observed for participants with *TP53* data, most notably in HIV status at enrollment. Among participants missing data for this SNP, 89% were HIV-uninfected at enrollment while there were no HIV-uninfected participants at baseline among those with *TP53* data. Hence, the exclusion due to lack of *TP53*

data resulted in an analytic sample entirely composed of participants who were HIV-infected at baseline. This study did not use a time-to-event analysis approach but rather one that treats person-time under observation equally, so the lack of observation of participants at and early after seroconversion does not threaten the validity of the analytic methods used.

Future work on this analysis will employ SNP imputation informed by 1000 Genomes or TopMed reference panel data to improve sample size. The imputation of *TP53* data will allow for inclusion of a substantial proportion of the study sample in analysis and the potential for subset analysis of participants who seroconverted under observation using time-to-event methods. This method of imputation would also be used to address any issues in missing SNP data that may arise in WIHS data. Furthermore, an ongoing genome-wide association study initiative will yield *TP53* SNP data on a large sample of MACS participants, allowing for the examination of these SNPs for both HIV-infected and uninfected participants.

Missing data in covariates is another limitation of this study which will be addressed in subsequent work. The study and results reported here largely used a last-observation-carried-forward imputation approach. In this approach, observations for missing time-varying covariates were carried forward for up to four visits representing two years. The strength of this approach is that the extent of missing data is reduced, however, it is also not without limitations. Among numerous limitations, this approach does not capture changes in covariates over time and makes no considerations about missing data mechanisms by assuming data are missing completely at random.

Future work on this analysis will use a multiple imputation approach to address missing data. This will first include considerations of missing data mechanisms, in part determining whether multiple imputation is appropriate or not, and building models to impute missing covariates where

appropriate. The benefits of multiple imputation compared to the approach used in this study are numerous, with preservation of the underlying distribution of observed covariates after imputation, and incorporation of missing data mechanisms and imputation based on observed covariates where appropriate, the most notable advantages.

Lastly, this study was limited by the method of adjustment for genetic ancestry. The study of genetic exposures requires adjustment for population stratification, a type of confounding in which cases and controls may have different underlying distributions of genetic ancestry. In this study, self-reported race was used to adjust for genetic ancestry, however, this was likely an inadequate, and in cases of misclassification inappropriate, means of adjustment. Future work will seek to improve adjustment for population stratification by using additional genetic data to compute genetic ancestry for each participant to be used to as the adjustment variable instead of self-reported race.

Upon addressing these limitations, this analysis will be able to examine differences in SNP effects by sex, address cancers which are only observed in women, and produce more precise estimates of SNP effects. Furthermore, the addition of WIHS data as well as SNP and covariate imputation will likely add sufficient events to enable multivariable modeling of adjusted SNP effects on incidence of NADCs and sensitivity analyses to assess the robustness of results, adding to the results reported here. Although it is unclear if an updated analysis incorporating these changes will estimate SNP effects that are statistically significant, it appears clear that these effects exist after adjustment for relevant covariates and that lack of statistical significance in the current study may be an issue of limited sample size.

There are also potential limitations which cannot be readily assessed or addressed in this study. These include limitations in ascertainment of cancers and differential loss to follow-up. The

ascertainment of cancers in this study was based on participant self-report with confirmatory medical record review as well as review of death certificates, and in some states, linkage to state cancer registries. Follow-up was stopped several years prior to when this data analysis was performed to allow for delays in cancer reporting and ascertainment. The potential limitations in cancer ascertainment are not due to these methods specifically but rather biases that exist in cancer screening, including lead-time bias and length-time bias, in which time or timing of cancer diagnoses may have differed by disease or frequency of screening, respectively.

Differential loss to follow-up was likely also an issue with participants with more advanced HIV disease missing study visits and being lost to follow-up. There were numerous cancer diagnoses recorded more than two years after participants' last visits that were excluded with participants censored at last follow-up. These participants may have had advanced disease and the exclusion of these cancer cases may have resulted in underestimates of associations of measures of disease severity and cancer outcomes. In the event that these limitations did impact the results of this study, it is likely that they biased results towards the null.

This study demonstrated a possible role for the p53-TLR8 axis of the immune response in the development of cancer in PLWHA, with participants carrying protective genotypes at both the *TP53* and *TLR8* SNPs at substantially reduced risk of cancer compared to those with neither protective genotype. The observation of this protective effect supports the hypothesis of higher affinity binding of p53 to its receptor element in *TLR8* and subsequent transcription of *TLR8*, resulting in a more robust immune response to viral pathogens.

The findings from this study identify relatively common SNPs with variable allele frequencies across ancestral populations which may help stratify PLWHA with respect to cancer risk. In spite of significant improvements in treatments for PLWHA, this population remains at elevated risk

for cancer compared to the general population. The introduction of HAART led to a significant reduction in ADCs, however, it is now well-established that PLWHA are also at higher risk for many NADCs than the general population. Among other cancer outcomes, this study specifically assessed the effects of SNPs that affect the immune response on the development of infection-related cancers, which includes both ADCs and a number of NADCs. Treatment with HAART prolongs survival with life expectancy approaching that of the general population with adequate access to care and full adherence, however, most do not regain full immune function and PLWHA are likely to remain at elevated risk for infection-related cancers. Therefore, these findings may be used to inform further developments in screening strategies, chemoprevention, and other interventions to reduce cancer risk in PLWHA.

Tables and Figures

Table 1: Baseline Characteristics (stratified by *TLR8* genotype)

Variable	Median (IQR) or n (%)	
	A* (N=538)	G* (N=222)
Age		
Median (IQR)	31.2 (26.8-36.5)	31.1 (26.9-35.8)
<35	366 (68.0)	158 (71.2)
35-44	138 (25.7)	56 (25.2)
45-54	28 (5.2)	8 (3.6)
55-64	4 (0.7)	0 (0.0)
≥65	2 (0.4)	0 (0.0)
Race		
White	446 (82.9)	176 (79.3)
Black	68 (12.6)	28 (12.6)
Latin/Hispanic	22 (4.1)	14 (6.3)
Other	2 (0.4)	4 (1.8)
TP53 genotype		
CC	84 (15.6)	31 (14.0)
CG	84 (15.6)	26 (11.7)
GG	25 (4.6)	13 (5.9)
Missing	345 (64.2)	152 (68.5)
BMI (kg/m^2)		
Median (IQR)	23.0 (21.6-24.8)	22.5 (21.1-24.4)
<25	406 (75.5)	174 (78.4)
25-30	101 (18.8)	36 (16.2)
>30	20 (3.7)	11 (5.0)
Missing	11 (2.0)	1 (0.5)
Alcohol consumption (drinks per week)		
0	24 (4.5)	17 (7.7)
1-3	151 (28.1)	55 (24.8)
4-13	244 (45.4)	101 (45.5)
>13	99 (18.4)	46 (20.7)
Missing	20 (3.7)	3 (1.4)
Smoking status		
Never	205 (38.1)	73 (32.9)
Former	97 (18.0)	35 (15.8)
Current	228 (42.4)	114 (51.4)
Missing	8 (1.5)	0 (0.0)
Cumulative pack-years smoking history		
Median (IQR)	2.7 (0.0-19.5)	3.7 (0.0-20.8)
<10.0	338 (62.8)	139 (62.6)
10.0-19.9	68 (12.6)	27 (12.2)
20.0-30.0	59 (11.0)	19 (8.6)
>30	66 (12.3)	37 (16.7)
Missing	7 (1.3)	0 (0.0)
HIV-infected		
No	304 (56.5)	138 (62.2)
Yes	235 (43.5)	84 (37.8)

Table 1: Baseline Characteristics (stratified by *TLR8* genotype; continued)

Variable	Median (IQR) or n (%)	
	A* (N=538)	G* (N=222)
CD4+ count (cells/mm^3)		
Median (IQR)	771 (559-1022)	791 (573-1010)
<200	4 (0.7)	4 (1.8)
200-349	21 (3.9)	5 (2.3)
350-500	66 (12.3)	25 (11.3)
>500	431 (80.1)	186 (83.8)
Missing	16 (3.0)	2 (0.9)
Viral load (copies/mm^3)		
Median (IQR)	0 (0-300)	0 (0-300)
<1000	291 (54.1)	134 (60.4)
≥1000	48 (8.9)	23 (10.4)
Missing	199 (37.0)	65 (29.3)

Abbreviations: BMI – body mass index, IQR – interquartile range; values reported are N (%) unless otherwise noted.

Table 2: Baseline Characteristics (stratified by *TP53* genotype)

Variable	Median (IQR) or n (%)			
	GG (N=115)	GC (N=110)	CC (N=38)	Not Tested (N=497)
Age				
Median (IQR)	31.1 (27.4-35.3)	33.0 (28.6-37.7)	32.2 (26.7-36.7)	30.8 (26.3-36.4)
<35	85 (73.9)	69 (62.7)	27 (71.1)	343 (69.0)
35-44	27 (23.5)	30 (27.3)	8 (21.1)	129 (26.0)
45-54	2 (1.7)	10 (9.1)	2 (5.3)	22 (4.4)
55-64	1 (0.9)	1 (0.9)	1 (2.6)	1 (0.2)
≥65	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.4)
Race				
White	98 (85.2)	80 (72.7)	14 (36.8)	430 (86.5)
Black	6 (5.2)	23 (20.9)	24 (63.2)	43 (8.7)
Latin/Hispanic	9 (7.8)	5 (4.5)	0 (0.0)	22 (4.4)
Other	2 (1.7)	2 (1.8)	0 (0.0)	2 (0.4)
TLR8				
G* genotype	31 (27.0)	26 (23.6)	13 (34.2)	152 (30.6)
A* genotype	84 (73.0)	84 (76.4)	25 (65.8)	345 (69.4)
BMI(kg/m^2)				
Median (IQR)	22.8 (21.6-24.2)	23.0 (21.6-24.5)	23.2 (21.7-25.7)	22.9 (21.4-24.9)
<25	97 (84.3)	88 (80.0)	26 (68.4)	369 (74.2)
25-30	14 (12.2)	18 (16.4)	10 (26.3)	95 (19.1)
>30	3 (2.6)	4 (3.6)	2 (5.3)	22 (4.4)
Missing	1 (0.9)	0 (0.0)	0 (0.0)	11 (2.2)
Alcohol consumption (drinks per week)				
0	5 (4.3)	9 (8.2)	5 (13.2)	22 (4.4)
1-3	35 (30.4)	30 (27.3)	10 (26.3)	131 (26.4)
4-13	50 (43.5)	50 (45.5)	11 (28.9)	234 (47.1)
>13	21 (18.3)	20 (18.2)	10 (26.3)	94 (18.9)
Missing	4 (3.5)	1 (0.9)	2 (5.3)	16 (3.2)
History of smoking				
Never	40 (34.8)	38 (34.5)	16 (42.1)	184 (37.0)
Former	19 (16.5)	22 (20.0)	7 (18.4)	84 (16.9)
Current	55 (47.8)	50 (45.5)	15 (39.5)	222 (44.7)
Missing	1 (0.9)	0 (0.0)	0 (0.0)	7 (1.4)
Cumulative pack-years of smoking history				
Median (IQR)	3.4 (0.0-19.5)	3.5 (0.0-21.9)	0.5 (0.0-12.0)	3.2 (0.0-18.8)
<10.0	72 (62.6)	66 (60.0)	28 (73.7)	311 (62.6)
10.0-19.9	14 (12.2)	15 (13.6)	2 (5.3)	64 (12.9)
20.0-30.0	10 (8.7)	14 (12.7)	5 (13.2)	49 (9.9)
>30.0	18 (15.7)	15 (13.6)	3 (7.9)	67 (13.5)
Missing	1 (0.9)	0 (0.0)	0 (0.0)	6 (1.2)
HIV-infected				
No	0 (0.0)	0 (0.0)	0 (0.0)	442 (88.8)
Yes	115 (100.0)	110 (100.0)	38 (100.0)	56 (11.2)

Table 2: Baseline Characteristics (stratified by *TP53* genotype; continued)

Variable	Median (IQR) or n (%)			
	GG (N=115)	GC (N=110)	CC (N=38)	Not Tested (N=497)
CD4+ count (cells/mm^3)				
Median (IQR)	651 (512-889)	604 (467-810)	577 (407-814)	848 (642-1097)
<200	2 (1.7)	3 (2.7)	2 (5.3)	1 (0.2)
200-349	4 (3.5)	10 (9.1)	0 (0.0)	12 (2.4)
350-500	18 (15.7)	23 (20.9)	12 (31.6)	38 (7.6)
>500	86 (74.8)	73 (66.4)	24 (63.2)	434 (87.3)
Missing	5 (4.3)	1 (0.9)	0 (0.0)	12 (2.4)
Viral load (copies/mm^3)				
Median (IQR)	27705 (8761-46935)	7440 (3763-27599)	20064 (4433-57680)	0 (0-40)
<1000	1 (0.9)	1 (0.9)	1 (2.6)	422 (84.9)
≥1000	15 (13.0)	16 (14.5)	8 (21.1)	32 (6.4)
Missing	99 (86.1)	93 (84.5)	29 (76.3)	43 (8.7)

Abbreviations: BMI – body mass index, IQR – interquartile range; values reported are N (%) unless otherwise noted.

Table 3: *TP53* and *TLR8* Allele Frequencies by Race

SNP	Self-reported race	Allele	Frequency
rs1042522, *TP53*	White	C	0.72
		G	0.28
	Black	C	0.33
		G	0.67
	Latin/Hispanic	C	0.82
		G	0.18
	Other	C	0.75
		G	0.25
rs3764880, *TLR8*	White	G	0.28
		A	0.72
	Black	G	0.29
		A	0.71
	Latin/Hispanic	G	0.39
		A	0.61
	Other	G	0.67
		A	0.33

Abbreviations: A – adenosine, C – cytosine, G – guanosine

Table 4: Diagnoses by Cancer Outcome

Diagnosis	SEER code	Any Cancer		AIDS-Defining Cancer		Non-AIDS-Defining Cancer		Infection-Related Cancer	
		Count	Percent	Count	Percent	Count	Percent	Count	Percent
Floor of mouth	20040	1	0.75	-	-	1	2.04	1	0.91
Tonsil	20070	1	0.75	-	-	1	2.04	1	0.91
Esophagus	21010	1	0.75	-	-	2	4.08	-	-
Hepatic flexure	21044	1	0.75	-	-	1	2.04	-	-
Large intestine, NOS	24049	1	0.75	-	-	1	2.04	-	-
Anus, anal canal and anorectum	21060	4	3.01	-	-	5	10.20	5	4.55
Liver	21071	3	2.26	-	-	3	6.12	3	2.73
Lung and bronchus	22030	3	2.26	-	-	3	6.12	3	2.73
Melanoma of the skin	25010	4	3.01	-	-	4	8.16	-	-
Prostate	28010	9	6.77	-	-	11	22.45	-	-
Testis	28020	2	1.50	-	-	2	4.08	-	-
Urinary bladder	29010	-	-	-	-	1	2.04	-	-
Kidney and renal pelvis	29020	2	1.50	-	-	2	4.08	-	-
Ureter	29030	1	0.75	-	-	1	2.04	-	-
Brain	31010	1	0.75	-	-	1	2.04	-	-
Thyroid	32010	3	2.26	-	-	3	6.12	-	-
Hodgkin lymphoma - nodal	33011	3	2.26	-	-	3	6.12	3	2.73
Non-Hodgkin lymphoma - nodal	33041	8	6.02	8	8.79	-	-	9	8.18
Non-Hodgkin lymphoma - extranodal	33042	11	8.27	11	12.09	-	-	13	11.82
Myeloma	34000	1	0.75	-	-	1	2.04	-	-
Kaposi sarcoma	36020	70	52.63	72	79.12	-	-	72	65.45
Miscellaneous	37000	3	2.26	-	-	3	6.12	-	-
Total		133	100.0	91	100.0	49	100.0	110	100.0

Table 5: Univariable Exact Poisson Regression – Any Cancer Diagnosis

Predictor	Events	Person-Years	IRR (95% CI)
Age			
<35	31	1961.36	1.00 (REF)
35-44	49	2927.18	1.04 (0.65-1.69)
45-54	29	2587.29	0.71 (0.41-1.22)
55-64	21	1041.29	1.28 (0.70-2.29)
≥65	3	171.88	1.10 (0.22-3.54)
<35	31	1961.36	1.00 (REF)
35-44	49	2927.18	1.04 (0.65-1.69)
45-54	29	2587.29	0.71 (0.41-1.22)
≥55	24	1213.16	1.25 (0.70-2.20)
TP53 genotype			
CC	18	1376.04	1.00 (REF)
CG	19	1323.02	1.10 (0.55-2.22)
GG	4	497.12	0.62 (0.15-1.87)
CC/CG	37	2699.06	1.00 (REF)
GG	4	497.12	0.59 (0.15-1.63)
Missing	92	5537.81	
TLR8 genotype			
G*	42	2425.88	1.00 (REF)
A*	91	6308.11	0.83 (0.57-1.23)
Race			
White	116	7205.89	1.00 (REF)
Black	12	959.32	0.78 (0.39-1.41)
Latin/Hispanic	4	484.62	0.51 (0.14-1.35)
Other	1	84.16	0.74 (0.02-4.19)
White	116	7205.89	1.00 (REF)
Black	12	959.32	0.78 (0.39-1.41)
Other	5	568.78	0.55 (0.17-1.31)
Non-White	17	1528.11	1.00 (REF)
White	116	7205.89	1.45 (0.87-2.57)
BMI (kg/m^2; lagged two years)			
<25	60	3085.49	1.00 (REF)
25-30	18	1505.84	0.61 (0.34-1.06)
>30	8	502.55	0.82 (0.34-1.72)
<25	60	3085.49	1.00 (REF)
≥25	26	2008.39	0.67 (0.40-1.07)
Missing	47	3640.12	
Alcohol consumption (drinks per week; lagged two years)			
0	11	810.50	1.00 (REF)
1-3	44	2336.75	1.39 (0.71-2.98)
4-13	27	1686.51	1.18 (0.57-2.64)
>13	8	438.11	1.35 (0.47-3.67)
0-3	55	3147.25	1.00 (REF)
>3	35	2124.62	0.94 (0.60-1.47)
Missing	43	3462.13	

Table 5: Univariable Exact Poisson Regression – Any Cancer Diagnosis (continued)

Predictor	Events	Person-Years	IRR (95% CI)
Smoking status (lagged two years)			
Never	29	1633.95	1.00 (REF)
Former	34	2080.94	0.92 (0.54-1.57)
Current	27	1554.28	0.98 (0.56-1.71)
Never	29	1633.95	1.00 (REF)
Former or Current	61	3635.22	0.95 (0.60-1.53)
Missing	43	3464.82	
Cumulative pack-years smoking history (lagged two years)			
<10.0	51	3147.89	1.00 (REF)
10.0-19.9	5	562.18	0.55 (0.17-1.37)
20.0-30.0	11	438.69	1.55 (0.73-3.01)
>30.0	20	1074.37	1.15 (0.65-1.96)
Missing	46	3510.90	
Smoking (lagged two years)			
Never	29	1633.95	1.00 (REF)
≤30 pack-year history	38	2514.81	0.85 (0.51-1.43)
>30 pack-year history	20	1074.37	1.05 (0.56-1.92)
Missing	46	5024.81	
CD4+ count (cells/mm^3; lagged two years)			
<200	15	332.79	1.00 (REF)
200-349	20	720.79	0.62 (0.30-1.29)
350-500	21	1128.71	0.41 (0.20-0.86)
>500	32	2990.27	0.24 (0.12-0.47)
<350	35	1053.58	1.00 (REF)
≥350	53	4118.98	0.39 (0.25-0.61)
Missing	45	3561.44	
Viral load (copies/mm^3; lagged two years)			
<1000	17	2341.57	1.00 (REF)
≥1000	62	2400.08	3.56 (2.05-6.49)
Missing	54	3992.35	
HAART (lagged two years)			
No	109	5709.30	1.00 (REF)
Yes	11	1704.63	0.34 (0.16-0.63)
Missing	13	1320.07	
HAART + Viral load (copies/mm^3; lagged two years)			
Not on HAART	109	5709.30	1.00 (REF)
On HAART, HIV viral load ≥1000	1	197.55	0.27 (0.01-1.51)
On HAART, HIV viral load <1000	10	1436.19	0.36 (0.17-0.90)
Missing	13	1390.96	
Treatment era			
1984-1995	89	4058.78	1.00 (REF)
1996-2014	44	4675.21	0.43 (0.29-0.62)
1984-1995	89	4058.78	1.00 (REF)
1996-2001	16	1630.01	0.45 (0.25-0.77)
2002-2014	28	3045.21	0.42 (0.26-0.65)

Abbreviations: BMI – body mass index, IRR – incidence rate ratio, HAART – highly-active antiretroviral therapy

Table 6: Univariable Exact Poisson Regression – AIDS-Defining Cancers

Predictor	Events	Person-Years	IRR (95% CI)
Age			
<35	29	1966.46	1.00 (REF)
35-44	41	2989.48	0.93 (0.56-1.55)
45-54	16	2640.42	0.41 (0.21-0.78)
55-64	7	1109.35	0.43 (0.16-1.00)
≥65	1	200.17	0.34 (0.01-2.04)
<35	29	1966.46	1.00 (REF)
35-44	41	2989.48	0.93 (0.56-1.55)
45-54	16	2640.42	0.41 (0.21-0.78)
≥55	8	1309.52	0.41 (0.16-0.93)
TP53 genotype			
CC	13	1416.56	1.00 (REF)
CG	15	1328.47	1.23 (0.55-2.81)
GG	3	498.28	0.66 (0.12-2.39)
CC/CG	28	2745.03	1.00 (REF)
GG	3	498.28	0.59 (0.11-1.91)
Missing	63	5662.56	
TLR8 genotype			
G*	30	2493.95	1.00 (REF)
A*	64	6411.92	0.83 (0.53-1.33)
Race			
White	81	7359.40	1.00 (REF)
Black	8	977.69	0.74 (0.31-1.54)
Latin/Hispanic	4	484.62	0.75 (0.20-2.00)
Other	1	84.16	1.08 (0.03-6.19)
White	81	7359.40	1.00 (REF)
Black	8	977.69	0.74 (0.31-1.54)
Other	5	568.78	0.80 (0.25-1.94)
Non-White	13	1546.47	1.00 (REF)
White	81	7359.40	1.31 (0.72-2.56)
BMI (kg/m^2; lagged two years)			
<25	42	3142.43	1.00 (REF)
25-30	14	1542.53	0.68 (0.34-1.27)
>30	3	522.59	0.43 (0.09-1.34)
<25	42	3142.43	1.00 (REF)
≥25	17	2065.12	0.62 (0.33-1.11)
Missing	35	3698.32	
Alcohol consumption (drinks per week; lagged two years)			
0	7	829.51	1.00 (REF)
1-3	29	2406.57	1.43 (0.61-3.86)
4-13	18	1715.39	1.24 (0.50-3.52)
>13	6	443.16	1.60 (0.45-5.58)
0-3	36	3236.08	1.00 (REF)
>3	24	2158.54	1.00 (0.57-1.72)
Missing	34	3511.25	

Table 6: Univariable Exact Poisson Regression – AIDS-Defining Cancers (continued)

Predictor	Events	Person-Years	IRR (95% CI)
Smoking status (lagged two years)			
Never	19	1668.86	1.00 (REF)
Former	21	2148.17	0.86 (0.44-1.69)
Current	20	1574.91	1.12 (0.57-2.21)
Never	19	1668.86	1.00 (REF)
Former or Current	41	3723.07	0.97 (0.55-1.76)
Missing	34	3513.94	
Cumulative pack-years smoking history (lagged two years)			
<10.0	35	3222.17	1.00 (REF)
10.0-19.9	3	569.78	0.48 (0.10-1.54)
20.0-30.0	7	442.36	1.46 (0.55-3.33)
>30.0	12	1111.58	0.99 (0.47-1.96)
Missing	37	3559.99	
Smoking (lagged two years)			
Never	19	1668.86	1.00 (REF)
≤30 pack-year history	26	2565.45	0.89 (0.47-1.70)
>30 pack-year history	12	1111.58	0.95 (0.42-2.06)
Missing	37	3559.99	
CD4+ count (cells/mm^3; lagged two years)			
<200	10	344.17	1.00 (REF)
200-349	15	734.61	0.70 (0.30-1.75)
350-500	15	1145.30	0.45 (0.19-1.12)
>500	19	3071.05	0.21 (0.09-0.51)
<350	25	1078.78	1.00 (REF)
≥350	34	4216.35	0.35 (0.20-0.61)
Missing	35	3610.74	
Viral load (copies/mm^3; lagged two years)			
<1000	3	2427.72	1.00 (REF)
≥1000	50	2430.63	16.65 (5.38-83.44)
Missing	41	4047.53	
HAART (lagged two years)			
No	87	5748.32	1.00 (REF)
Yes	0	1791.91	0.03 (undefined-0.14)
Missing	7	1365.65	
HAART + Viral load (copies/mm^3; lagged two years)			
Not on HAART	87	5748.32	1.00 (REF)
On HAART, HIV viral load ≥1000	0	205.08	0.22 (0.00-1.21)
On HAART, HIV viral load <1000	0	1514.28	0.03 (0.00-0.16)
Missing	7	1438.20	
Treatment era			
1984-1995	80	4074.69	1.00 (REF)
1996-2014	14	4831.19	0.15 (0.08-0.26)
1984-1995	80	4074.69	1.00 (REF)
1996-2001	8	1654.77	0.25 (0.10-0.51)
2002-2014	6	3176.41	0.10 (0.03-0.22)

Abbreviations: BMI – body mass index, IRR – incidence rate ratio, HAART – highly-active antiretroviral therapy

Table 7: Univariable Exact Poisson Regression – Non-AIDS-Defining Cancers

Predictor	Events	Person-Years	IRR (95% CI)
Age			
<35	3	1988.60	1.00 (REF)
35-44	11	3032.59	2.40 (0.64-13.42)
45-54	16	2666.69	3.98 (1.14-21.30)
55-64	15	1084.16	9.17 (2.59-49.42)
≥65	4	177.62	14.93 (2.53-101.91)
<35	3	1988.60	1.00 (REF)
35-44	11	3032.59	2.40 (0.64-13.42)
45-54	16	2666.69	3.98 (1.14-21.30)
≥55	19	1261.78	9.98 (2.94-52.67)
TP53 genotype			
CC	6	1416.29	1.00 (REF)
CG	6	1336.06	1.06 (0.28-3.97)
GG	1	495.43	0.48 (0.01-3.93)
CC/CG	12	2752.35	1.00 (REF)
GG	1	495.43	0.46 (0.01-3.13)
Missing	36	5701.88	
TLR8 genotype			
G*	16	2483.99	1.00 (REF)
A*	33	6465.67	0.79 (0.42-1.54)
Race			
White	44	7424.10	1.00 (REF)
Black	5	959.79	0.88 (0.27-2.21)
Latin/Hispanic	0	481.55	0.24 (0.00-1.35)
Other	0	84.22	1.40 (0.00-7.71)
White	44	7424.10	1.00 (REF)
Black	5	959.79	0.88 (0.27-2.21)
Other	0	565.77	0.21 (0.00-1.15)
Non-White	5	959.79	1.00 (REF)
White	44	7424.10	1.81 (0.72-5.84)
BMI (kg/m^2; lagged two years)			
<25	23	3085.49	1.00 (REF)
25-30	5	1505.84	0.45 (0.13-1.21)
>30	5	502.55	1.37 (0.41-3.68)
<25	23	3085.49	1.00 (REF)
≥25	10	2046.41	0.68 (0.29-1.47)
Missing	16	3640.12	
Alcohol consumption (drinks per week; lagged two years)			
0	6	833.72	1.00 (REF)
1-3	17	2421.62	0.98 (0.37-3.02)
4-13	12	1722.80	0.97 (0.34-3.14)
>13	2	439.04	0.63 (0.06-3.54)
0-3	23	3255.34	1.00 (REF)
>3	14	2161.84	0.92 (0.44-1.86)
Missing	12	3532.48	

Table 7: Univariable Exact Poisson Regression – Non-AIDS-Defining Cancers (continued)

Predictor	Events	Person-Years	IRR (95% CI)
Smoking status (lagged two years)			
Never	11	1694.40	1.00 (REF)
Former	17	2147.42	1.22 (0.54-2.88)
Current	9	1572.67	0.88 (0.32-2.34)
Never	11	1694.40	1.00 (REF)
Former or Current	26	3720.09	1.08 (0.51-2.41)
Missing	12	3535.17	
Cumulative pack-years smoking history (lagged two years)			
<10.0	19	3261.16	1.00 (REF)
10.0-19.9	2	564.26	0.61 (0.07-2.52)
20.0-30.0	5	455.53	1.88 (0.55-5.22)
>30.0	11	1083.15	1.74 (0.75-3.85)
Missing	12	3585.57	
Smoking (lagged two years)			
Never	11	1694.40	1.00 (REF)
≤30 pack-year history	15	2584.44	0.89 (0.38-2.15)
>30 pack-year history	11	1083.15	1.56 (0.62-3.98)
Missing	22	3587.68	
CD4+ count (cells/mm^3; lagged two years)			
<200	7	375.90	1.00 (REF)
200-349	5	755.59	0.36 (0.09-1.30)
350-500	7	1161.34	0.32 (0.10-1.08)
>500	17	3018.34	0.30 (0.12-0.86)
<350	12	1131.49	1.00 (REF)
≥350	24	4179.68	0.54 (0.26-1.19)
Missing	13	3638.49	
Viral load (copies/mm^3; lagged two years)			
<1000	16	2392.22	1.00 (REF)
≥1000	17	2476.20	1.03 (0.49-2.17)
Missing	16	4081.24	
HAART (lagged two years)			
No	28	5809.07	1.00 (REF)
Yes	14	1767.45	1.64 (0.80-3.23)
Missing	7	1373.15	
HAART + Viral load (copies/mm^3; lagged two years)			
Not on HAART	28	5809.07	1.00 (REF)
On HAART, HIV viral load ≥1000	2	208.27	1.99 (0.23-7.90)
On HAART, HIV viral load <1000	12	1482.14	1.68 (0.78-3.41)
Missing	7	1450.19	
Treatment era			
1984-1995	13	4116.44	1.00 (REF)
1996-2014	36	4833.22	2.36 (1.22-4.85)
1984-1995	13	4116.44	1.00 (REF)
1996-2001	11	1697.46	2.05 (0.83-4.96)
2002-2014	25	3135.76	2.52 (1.24-5.37)

Abbreviations: BMI – body mass index, IRR – incidence rate ratio, HAART – highly-active antiretroviral therapy

Table 8: Univariable Exact Poisson Regression – Infection-Related Cancers

Predictor	Events	Person-Years	IRR (95% CI)
Age			
<35	30	1964.85	1.00 (REF)
35-44	45	2984.13	0.99 (0.61-1.62)
45-54	26	2611.13	0.65 (0.37-1.14)
55-64	8	1089.73	0.48 (0.19-1.07)
≥65	1	197.79	0.33 (0.01-1.99)
<35	30	1964.85	1.00 (REF)
35-44	45	2984.13	0.99 (0.61-1.62)
45-54	26	2611.13	0.65 (0.37-1.14)
≥55	9	1287.53	0.46 (0.19-0.99)
TP53 genotype			
CC	16	1396.40	1.00 (REF)
CG	19	1322.47	1.25 (0.61-2.61)
GG	3	498.28	0.53 (0.10-1.84)
CC/CG	35	2718.88	1.00 (REF)
GG	3	498.28	0.47 (0.09-1.48)
Missing	72	5630.47	
TLR8 genotype			
G*	34	2482.75	1.00 (REF)
A*	76	6364.89	0.87 (0.57-1.35)
Race			
White	96	7313.58	1.00 (REF)
Black	9	965.27	0.71 (0.32-1.41)
Latin/Hispanic	4	484.62	0.63 (0.17-1.66)
Other	1	84.16	0.91 (0.02-5.17)
White	96	7313.58	1.00 (REF)
Black	9	965.27	0.71 (0.32-1.41)
Other	5	568.78	0.67 (0.21-1.62)
Non-White	14	1534.05	1.00 (REF)
White	96	7313.58	1.44 (0.82-2.73)
BMI (kg/m^2; lagged two years)			
<25	51	3114.82	1.00 (REF)
25-30	14	1542.03	0.55 (0.28-1.02)
>30	6	507.14	0.72 (0.25-1.68)
<25	51	3114.82	1.00 (REF)
>25	20	2049.17	0.60 (0.34-1.02)
Missing	39	3683.65	
Alcohol consumption (drinks per week; lagged two years)			
0	8	827.02	1.00 (REF)
1-3	35	2387.25	1.52 (0.39-3.78)
4-13	23	1697.22	1.40 (0.60-3.62)
>13	7	437.81	1.65 (0.51-5.22)
0-3	43	3214.28	1.00 (REF)
>3	30	2135.03	1.05 (0.64-1.71)
Missing	37	3498.33	

Table 8: Univariable Exact Poisson Regression – Infection-Related Cancers (continued)

Predictor	Events	Person-Years	IRR (95% CI)
Smoking status (lagged two years)			
Never	22	1653.48	1.00 (REF)
Former	26	2137.43	0.91 (0.50-1.69)
Current	25	1555.70	1.21 (0.65-2.25)
Never	22	1653.48	1.00 (REF)
Former or Current	51	3693.13	1.04 (0.62-1.80)
Missing	37	3501.02	
Cumulative pack-years smoking history (lagged two years)			
<10.0	40	3202.62	1.00 (REF)
10.0-19.9	5	562.18	0.71 (0.22-1.80)
20.0-30.0	10	438.39	1.83 (0.81-3.72)
>30.0	15	1097.37	1.09 (0.56-2.02)
Missing	40	3547.07	
Smoking (lagged two years)			
Never	22	1653.48	1.00 (REF)
≤30 pack-year history	15	1000.57	1.13 (0.54-2.27)
>30 pack-year history	15	1097.37	1.03 (0.50-2.07)
Missing	58	5096.22	
CD4+ count (cells/mm^3; lagged two years)			
<200	16	335.65	1.00 (REF)
200-349	16	725.67	0.46 (0.22-0.99)
350-500	17	1140.67	0.31 (0.15-0.66)
>500	23	3047.82	0.16 (0.08-0.32)
<350	32	1061.32	1.00 (REF)
≥350	40	4188.50	0.32 (0.19-0.52)
Missing	38	3597.82	
Viral load (copies/mm^3; lagged two years)			
<1000	5	2396.71	1.00 (REF)
≥1000	59	2417.71	11.70 (4.74-37.35)
Missing	46	4033.21	
HAART (lagged two years)			
No	100	5734.40	1.00 (REF)
Yes	2	1761.12	0.07 (0.01-0.24)
Missing	8	1352.12	
HAART + Viral load (copies/mm^3; lagged two years)			
Not on HAART	100	5734.40	1.00 (REF)
On HAART, HIV viral load ≥1000	1	201.43	0.28 (0.01-1.62)
On HAART, HIV viral load <1000	1	1487.13	0.04 (0.01-0.22)
Missing	8	1424.68	
Treatment era			
1984-1995	85	4070.43	1.00 (REF)
1996-2014	25	4777.21	0.25 (0.15-0.40)
1984-1995	85	4070.43	1.00 (REF)
1996-2001	14	1645.39	0.41 (0.21-0.72)
2002-2014	11	3131.82	0.17 (0.08-0.32)

Abbreviations: BMI – body mass index, IRR – incidence rate ratio, HAART – highly-active antiretroviral therapy

Table 9: Multivariable Exact Poisson Regression – All Outcomes

Predictor	Any Cancer IRR (95% CI)	AIDS-defining Cancer IRR (95% CI)	Infection-Related Cancer IRR (95% CI)
TP53 genotype			
CC, CG	1.00 (REF)	1.00 (REF)	1.00 (REF)
GG	0.65 (0.01-5.74)	1.01 (0.02-10.74)	0.86 (0.02-8.00)
TLR8 genotype			
G*	1.00 (REF)	1.00 (REF)	1.00 (REF)
A*	0.92 (0.33-2.93)	0.84 (0.22-3.90)	1.01 (0.33-3.69)
Joint Effect (*TP53*GG**TLR8* A*)			
No	1.00 (REF)	1.00 (REF)	1.00 (REF)
Yes	0.46 (0.01-42.71)	0.47 (0.01-48.55)	0.40 (0.01-39.37)
Race			
White	1.00 (REF)	1.00 (REF)	1.00 (REF)
Black	0.99 (0.16-3.82)	1.79 (0.27-7.87)	1.19 (0.19-4.72)
Other	0.80 (0.02-5.27)	1.52 (0.03-11.59)	1.09 (0.03-7.53)
BMI (kg/m^2; lagged two years)			
Underweight/Normal	1.00 (REF)	1.00 (REF)	1.00 (REF)
Overweight/Obese	0.48 (0.12-1.44)	0.19 (0.01-1.35)	0.24 (0.03-1.01)
Smoking history (lagged two years)			
Never	1.00 (REF)	1.00 (REF)	1.00 (REF)
≤30 pack-year history	0.78 (0.30-2.12)	0.74 (0.18-3.11)	0.95 (0.33-2.91)
>30 pack-year history	0.39 (0.07-1.60)	0.29 (0.01-2.60)	0.53 (0.09-2.37)
CD4+ Count (cells/mm^3; lagged two years)			
<350	1.00 (REF)	1.00 (REF)	1.00 (REF)
≥350	0.26 (0.11-0.64)	0.36 (0.09-1.32)	0.26 (0.09-0.68)
HAART + Viral load (copies/mm^3; lagged two years)			
Not on HAART	1.00 (REF)	1.00 (REF)	1.00 (REF)
On HAART, viral load ≥1000	0.78 (0.00-4.65)	1.14 (0.00-7.42)	1.25 (0.03-8.56)
On HAART, viral load <1000	0.49 (0.05-2.02)	0.22 (0.00-1.35)	0.19 (0.00-1.10)
Alcohol consumption (drinks per week; lagged two years)			
0-3	N/I	N/I	1.00 (REF)
>3	N/I	N/I	1.50 (0.54-4.10)

Abbreviations: BMI – body mass index, IRR – incidence rate ratio, HAART – highly-active antiretroviral therapy, N/I – not included

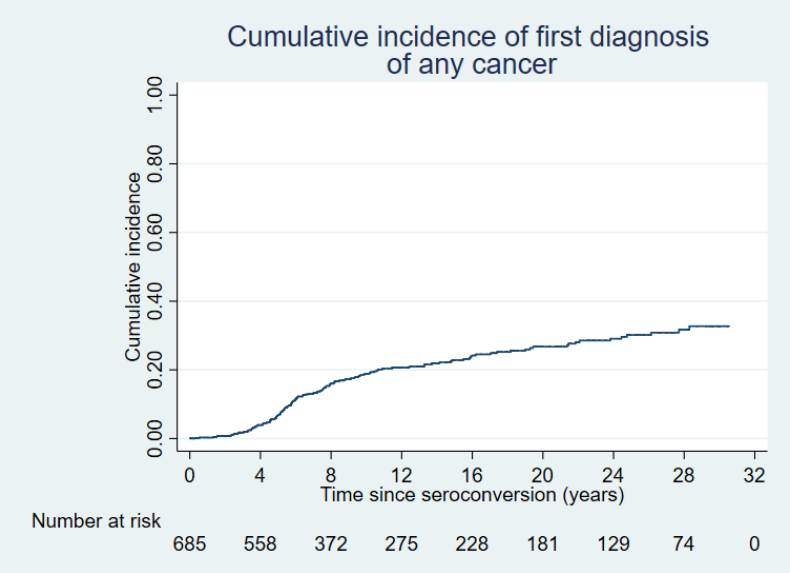

Figure 1: Cumulative Incidence of Any Cancer Diagnosis

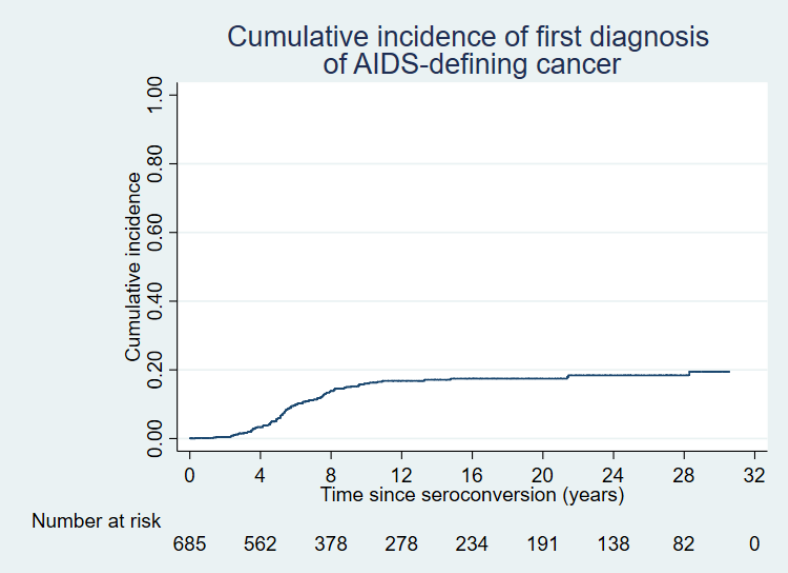

Figure 2: Cumulative Incidence of AIDS-Defining Cancers

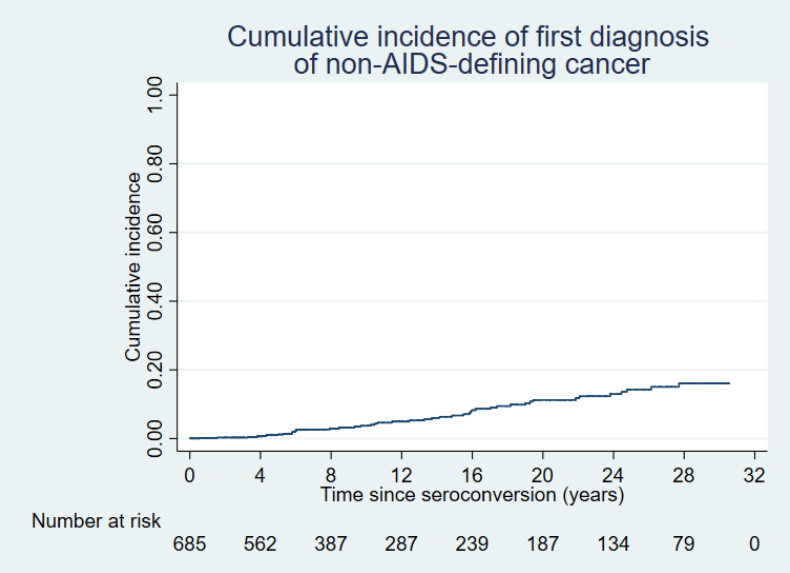

Figure 3: Cumulative Incidence of Non-AIDS-Defining Cancers

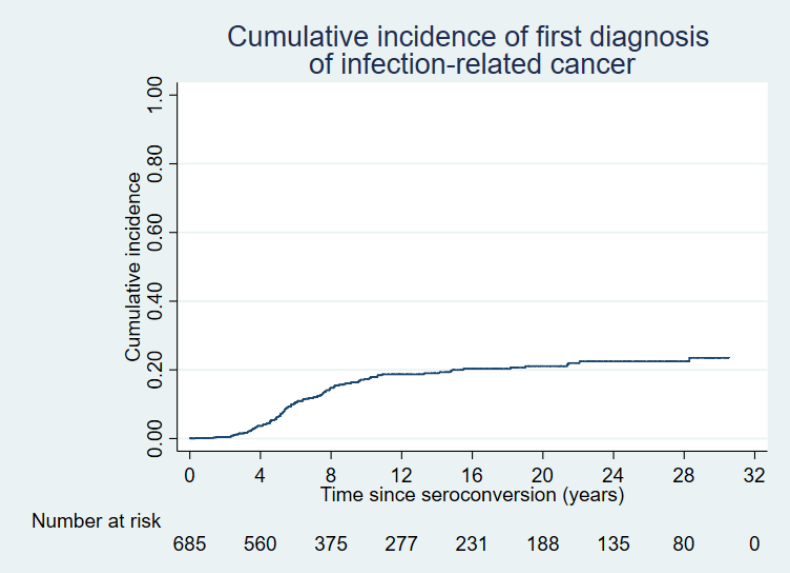

Figure 4: Cumulative Incidence of Infection-Related Cancers

Study Acknowledgements

Data in this manuscript were collected by the Multicenter AIDS Cohort Study (MACS). MACS (Principal Investigators): Johns Hopkins University Bloomberg School of Public Health (Joseph Margolick, Todd Brown), U01-AI35042; Northwestern University (Steven Wolinsky), U01-AI35039; University of California, Los Angeles (Roger Detels, Otoniel Martinez-Maza, Otto Yang), U01-AI35040; University of Pittsburgh (Charles Rinaldo, Lawrence Kingsley, Jeremy Martinson), U01-AI35041; the Center for Analysis and Management of MACS, Johns Hopkins University Bloomberg School of Public Health (Lisa Jacobson, Gypsyamber D'Souza), UM1-AI35043. The MACS is funded primarily by the National Institute of Allergy and Infectious Diseases (NIAID), with additional co-funding from the National Cancer Institute (NCI), the National Institute on Drug Abuse (NIDA), and the National Institute of Mental Health (NIMH). Targeted supplemental funding for specific projects was also provided by the National Heart, Lung, and Blood Institute (NHLBI), and the National Institute on Deafness and Communication Disorders (NIDCD). MACS data collection is also supported by UL1-TR001079 (JHU ICTR) from the National Center for Advancing Translational Sciences (NCATS) a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research. The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health (NIH), Johns Hopkins ICTR, or NCATS. The MACS website is located at http://aidscohortstudy.org/.

Cancer incidence data were provided by the following state agencies: 1) Maryland Cancer Registry, Center for Cancer Prevention and Control, Department of Health and Mental Hygiene, Baltimore, MD 21201; 2) Illinois Department of Public Health, Illinois State Cancer Registry; 3) Bureau of Health Statistics & Research, Pennsylvania Department of Health, Harrisburg,

Pennsylvania; 4) Ohio Cancer Incidence Surveillance System (OCISS), Ohio Department of Health (ODH), a cancer registry partially supported in the National Program of Cancer Registries at the Centers for Disease Control and Prevention (CDC) through Cooperative Agreement #5U58DP000795-05; and 5) California Department of Public Health pursuant to California Health and Safety Code Section 103885; CDC's National Program of Cancer Registries, under cooperative agreement 5NU58DP003862-04/DP003862; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute. We acknowledge the State of Maryland, the Maryland Cigarette Restitution Fund, and the National Program of Cancer Registries of the CDC for the funds that support the collection and availability of the cancer registry data. The analyses, findings, interpretations and conclusions of this report are those of the authors. No endorsement by any of the states providing data, the National Cancer Institute, the CDC or their Contractors and Subcontractors is intended nor should be inferred.

References

1. Centers for Disease Control. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men—New York City and California. MMWR Morb Mortal Wkly Repo. 1981;30(25):305-308
2. Barre-Sinoussi F, Chermann JC, Rey F, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk of acquired immune deficiency syndrome (AIDS). Science. 1983;220(4599):868-871
3. Gallo RC, Salahuddin SZ, Popovic M, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science. 1984;224(4648):500-503
4. Popovic M, Sarngadharan MG, Read E, Gallo RC. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science. 1984;224(4648):497-500
5. Levy JA, Hoffman AD, Kramer SM, Landis JA, Shimabukuro JM, Oshiro LS. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science. 1984;225(4664):840-842
6. Turner BG, Summers MF. Structural biology of HIV. J Mol Bio. 1999;285(1):1-32
7. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature. 1984;312(5996):763-767
8. Klatzmann D, Champagne E, Chamaret S, et al. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. Nature. 1984;312(5996):767-768
9. Deng H, Liu R, Ellmeier W, et al. Identification of a major co-receptor for primary isolates of HIV-1. Nature. 1996;381(6584):664-666
10. Dragic T, Litwin V, Allaway GP, et al. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature. 1996;381(6584):667-673
11. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science. 1996;272(5263):872-877
12. Shaw GM, Hunter E. HIV Transmission. Cold Spring Harb Perspect Med. 2012;2(11):a006965
13. GBD 2015 HIV Collaborators. Estimates of global, regional, and national incidence, prevalence and mortality of HIV, 1980-2015: the Global Burden of Disease Study 2015. Lancet HIV. 2016;3(8):e361-e387
14. Maartens G, Celum C, Lewin SR. HIV infection: epidemiology, pathogenesis, treatment, and prevention. Lancet. 2014;384(9939):258-271
15. Greene WC. A history of AIDS: looking back to see ahead. Eur J Immunol. 2007;37 Suppl 1:S94-S102
16. Barre-Sinoussi F, Ross AL, Delfraissy JF. Past, present and future: 30 years of HIV research. Nat Rev Microbiol. 2013;11(12):877-883
17. Fischl MA, Richman DD, Grieco MH, et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. N Engl J Med. 1987;317(4):185-191
18. Larder BA, Kemp SD. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). Science. 1989;246(4934):1155-1158

19. Collier AC, Coombs RW, Schoenfeld DA, et al. Treatment of human immunodeficiency virus infection with saquinovir, zidovudine, and zalcitabine. AIDS Clinical Trials Group. N Engl J Med. 1996;334(16):1011-1017
20. D'Aquila RT, Hughes MD, Johnson VA, et al. Nevirapine, zidovudine, and didanosine compared with zidovudine and didanosine in patients with HIV-1 infection. A randomized, double-blind, placebo-controlled trial. National Institute of Allergy and Infectious Diseases AIDS Clinical Trials Group Protocol 241 Investigators. Ann Intern Med. 1996;124(12):1019-1030
21. Hammer SM, Squires KE, Hughes MD, et al. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. N Engl J Med. 1997;337(11):725-733
22. Steigbigel RT, Cooper DA, Kumar PN, et al. Raltegravir with optimized background therapy for resistant HIV-1 infection. N Engl J Med. 2008;359(4):339-354
23. UN Joint Programme on HIV/AIDS (UNAIDS), UNAIDS Data 2018. Available at: http://www.unaids.org/sites/default/files/media_asset/unaids-data-2018_en.pdf
24. Weber R, Ruppik M, Rickenbach M, et al. Decreasing mortality and changing patterns of causes of death in the Swiss HIV Cohort Study. HIV Med. 2013;14(4):195-207
25. Bhaskaran K, Hamouda O, Sannes M, et al. Changes in the risk of death after HIV seroconversion compared with mortality in the general population. JAMA. 2008;300(1):51-59
26. Deeks SG, Lewin SR, Havlir DV. The End of AIDS: HIV Infection as a Chronic Disease. Lancet. 2013;382(9903):1525-1533
27. Trickey A, May MT, Vehreschild JJ, et al. Survival of HIV-positive patients starting antiretroviral therapy between 1996 and 2013: a collaborative analysis of cohort studies. Lancet HIV. 2017;4(8):e349-e356
28. Althoff KN, McGinnis KA, Wyatt CM, et al. Comparison of risk and age at diagnosis of myocardial infarction, end-stage renal disease, and non-AIDS-defining cancer in HIV-infected versus uninfected adults. Clin Infect Dis. 2015;60(4):627-638
29. Wong C, Gange SJ, Moore RD, et al. Multimorbidity Among Persons Living with Human Immunodeficiency Virus in the United States. Clin Infect Dis. 2018;66(8):1230-1238
30. Freiberg MS, Chang CC, Kuller LH, et al. HIV infection and the risk of acute myocardial infarction. JAMA Intern Med. 2013;173(8):614-622
31. Saylor D, Dickens AM, Sacktor N, et al. HIV-associated neurocognitive disorder – pathogenesis and prospects for treatment. Nat Rev Neurol. 2016;12(4):234-248
32. Robbins HA, Pfeiffer RM, Shiels MS, Li J, Hall HI, Engels EA. Excess cancers among HIV-infected people in the United States. J Natl Cancer Inst. 2015;107(4):dju503
33. Patel P, Hanson DL, Sullivan PS, et al. Incidence of types of cancer among HIV-infected persons compared with the general population in the United States, 1992-2003. Ann Intern Med. 2008;148(10):728-736
34. Engels EA, Yanik EL, Wheeler W, et al. Cancer-Attributable Mortality Among People With Treated Human Immunodeficiency Virus Infection in North America. Clin Infect Dis. 2017;65(4):636-64

35. Morlat P, Roussillon C, Henard S, et al. Causes of death among HIV-infected patients in France in 2010 (national survey): trends since 2000. AIDS. 2014;28(8):1181-1191
36. Goehringer F, Bonnet F, Salmon D, et al. Causes of Death in HIV-Infected Individuals with Immunovirologic Success in a National Prospective Survey. AIDS Res Hum Retroviruses. 2017;33(2):187-193
37. Shiels MS, Pfeiffer RM, Gail MH, et al. Cancer burden in the HIV-infected population in the United States. J Natl Cancer Inst. 2011;103(9):753-762
38. Hessol NA, Seaberg EC, Preston-Martin S, et al. Cancer risk among participants in the women's interagency HIV study. J Acquir Immune Defic Syndr. 2004;36(4):978-985
39. Mbulawa ZZ, Marais DJ, Johnson LF, Coetzee D, Williamson AL. Impact of human immunodeficiency virus on the natural history of human papillomavirus genital infection in South African men and women. J Infect Dis. 2012;206(1):15-27
40. Palefsky JM, Minkoff H, Kalish LA, et al. Cervicovaginal human papillomavirus infection in human immunodeficiency virus-1 (HIV)-positive and high-risk HIV-negative women. J Natl Cancer Inst. 1999;91(3):226-236
41. Goncalves PH, Uldrick TS, Yarchoan R. HIV-associated Kaposi sarcoma and related diseases. AIDS. 2017;31(14):1903-1916
42. Reekie J, Kosa C, Engsig F, et al. Relationship between current level of immunodeficiency and non-acquired immunodeficiency syndrome-defining malignancies. Cancer. 2010;116(22):5306-5315
43. Kesselring A, Gras L, Smit C, et al. Immunodeficiency as a risk factor for non-AIDS-defining malignancies in HIV-1-infected patients receiving combination antiretroviral therapy. Clin Infect Dis. 2011;52(12):1458-1465
44. Guiguet M, Boue F, Cadranel J, et al. Effect of immunodeficiency, HIV viral load, and antiretroviral therapy on the risk of individual malignancies (FHDH-ANRS CO4): a prospective cohort study. Lancet Oncol. 2009;10(12):1152-1159
45. Mdodo R, Frazier EL, Dube SR, et al. Cigarette smoking prevalence among adults with HIV compared with the general adult population in the United States: cross-sectional surveys. Ann Intern Med. 2015;162(5):335-344
46. Galvan FH, Bing EG, Fleishman JA, et al. The prevalence of alcohol consumption and heavy drinking among people with HIV in the United States: results from the HIV Cost and Services Utilization Study. J Stud Alcohol. 2002;63(2):179-186
47. Marhsall BDL, Tate JP, McGinnis KA, et al. Long-term alcohol use patterns and HIV disease severity. AIDS. 2017;31(9):1313-1321
48. Bonfanti P, Giannattasio C, Ricci E, et al. HIV and metabolic syndrome: a comparison with the general population. J Acquir Immune Defic Syndr. 2007;45(4):426-431
49. Sobieszczyk ME, Hoover DR, Anastos K, et al. Prevalence and predictors of metabolic syndrome among HIV-infected and HIV-uninfected women in the Women's Interagency HIV Study. J Acquir Immune Defic Syndr. 2008;48(3):272-280
50. Mulligan K, Grunfeld C, Tai VW, et al. Hyperlipidemia and insulin resistance are induced by protease inhibitors independent of changes in body composition in patients with HIV infection. J Acquir Immune Defic Syndr. 2000;23(1):35-43

51. Riddler SA, Smit E, Cole SR, et al. Impact of HIV infection and HAART on serum lipids in men. JAMA. 2003;289(22):2978-2982
52. Crum-Cianflone N, Hullsiek KH, Marconi V, et al. Trends in the incidence of cancers among HIV-infected persons and the impact of antiretroviral therapy: a 20-year cohort study. AIDS. 2009;23(1):41-50
53. Centers for Disease Control. Revision of the case definition of acquired immunodeficiency syndrome for national reporting – United States. MMWR Morb Mortal Wkly Rep. 1985;34(25):373-375
54. Centers for Disease Control and Prevention. Impact of the expanded AIDS surveillance case definition on AIDS case reporting – United States, first quarter, 1993. MMWR Morb Mortal Wkly Rep. 1993;42(16):308-310
55. Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical cancer. Lancet. 2019;393(10167):169-182
56. Cesarman E, Chang Y, Moore PS, Said JW, Knowles DM. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in AIDS-related body-cavity-based lymphomas. N Engl J Med. 1995;332(18):1186-1191
57. Guillet S, Gerard L, Meignin V, et al. Classic and extracavitary primary effusion lymphoma in 51 HIV-infected patients from a single institution. Am J Hematol. 2016;92(2):233-237
58. Saha A, Robertson ES. Epstein-Barr virus-associated B-cell lymphomas: pathogenesis and clinical outcomes. Clin Cancer Res. 2011;17(10):3056-3063
59. Vockerodt M, Yap LF, Shannon-Lowe C, et al. The Epstein-Barr virus and the pathogenesis of lymphoma. J Pathol. 2015;235(2):312-322
60. de Sanjose S, Benavente Y, Vajdic CM, et al. Hepatitis C and non-Hodgkin lymphoma among 4784 cases and 6269 controls from the International Lymphoma Epidemiology Consortium. Clin Gastroenterol Hepatol. 2008;6(4):451-458
61. Giordano TP, Henderson L, Landgren O, et al. Risk of non-Hodgkin lymphoma and lymphoproliferative precursor disease in US veterans with hepatitis C virus. JAMA. 2007;297(18):2010-2017
62. Armitage JO, Gascoyne RD, Lunning MA, Cavalli F. Non-Hodgkin lymphoma. Lancet. 2017;390(10091):298-310
63. Morton LM, Slager SL, Cerhan JR, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes: the InterLymph Non-Hodgkin Lymphoma Subtypes Project. J Natl Cancer Inst Monogr. 2014;2014(48):130-144
64. Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol. 2011;29(32):4294-4301
65. de Martel C, Plummer M, Vignat J, Franceschi S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. Int J Cancer. 2017;141(4):664-670
66. Carbone A, Gloghini A, Caruso A, De Paoli P, Dolcetti R. The impact of EBV and HIV infection on the microenvironmental niche underlying Hodgkin lymphoma pathogenesis. Int J Cancer. 2017;140(6):1233-1245
67. Ringhelhan M, O'Connor T, Protzer U, Heikenwalder M. The direct and indirect roles of HBV in liver cancer: prospective markers for HCC screening and potential therapeutic targets. J Pathol. 2015;235(2):355-367

68. Hessol NA, Martinez-Maza O, Levine AM, et al. Lung cancer incidence and survival among HIV-infected and uninfected women and men. AIDS. 2015;29(10):1183-1193
69. Blaser MJ, Perez-Perez GI, Kleanthous H, et al. Infection with Helicobacter pylori strains possession cagA is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res. 1995;55(10):2111-2115
70. Coghille AE, Engels EA, Schymura MJ, Mahale P, Shiels MS. Risk of Breast, Prostate, and Colorectal Cancer Diagnoses Among HIV-Infected Individuals in the United States. J Natl Cancer Inst. 2018;110(9):959-966
71. Goedert JJ, Schairer C, McNeel TS, et al. Risk of breast, ovary, and uterine corpus cancers among 85,268 women with AIDS. Br J Cancer. 2006;95(5):642-648
72. Shiels MS, Goedert JJ, Moore RD, Platz EA, Engels EA. Reduced risk of prostate cancer in U.S. Men with AIDS. Cancer Epidemiol Biomarkers Prev. 2010;19(11):2910-2915
73. Marcus JL, Chao CR, Leyden WA, et al. Prostate cancer incidence and prostate-specific antigen testing among HIV-positive and HIV-negative men. J Acquir Immune Defic Syndr. 2014;66(5):495-502
74. Hessol NA, Napolitano LA, Smith D, et al. HIV tropism and decreased risk of breast cancer. PLoS One. 2010;5(12):e14349
75. de Martel C, Shiels MS, Franceschi S, et al. Cancers attributable to infections among adults with HIV in the United States. AIDS. 2015;29(16):2173-2181
76. Silverberg MJ, Chao C, Leyden WA, et al. HIV infection and the risk of cancers with and without a known infectious cause. AIDS. 2009;23(17):2337-2345
77. Hernandez-Ramirez RU, Shiels MS, Dubrow R, Engels EA. Cancer risk in HIV-infected people in the USA from 1996 to 2012: a population-based, registry-linkage study. Lancet HIV. 2017;4(11):e495-e504
78. Engels EA, Biggar RJ, Hall HI, et al. Cancer risk in people infected with human immunodeficiency virus in the United States. Int J Cancer. 2008;123(1):187-194
79. Franceschi S, Maso LD, Rickenbach M, et al. Kaposi sarcoma incidence in the Swiss HIV Cohort Study before and after highly active antiretroviral therapy. Br J Cancer. 2008;99(5):800-804
80. Polesel J, Clifford GM, Rickenbach M, et al. Non-Hodgkin lymphoma incidence in the Swiss HIV Cohort Study before and after highly active antiretroviral therapy. AIDS. 2008;22(2):301-306
81. Autran B, Carcelain G, Li TS, et al. Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease. Science. 1997;277(5322):112-116
82. Moore RD, Keruly JC. CD4+ cell count 6 years after commencement of highly active antiretroviral therapy in persons with sustained virologic suppression. Clin Infect Dis. 2007;44(3):441-446
83. Robbins GK, Spritzler JG, Chan ES, et al. Incomplete reconstitution of T cell subsets on combination antiretroviral therapy in the AIDS Clinical Trials Group protocol 384. Clin Infect Dis. 2009;48(3):350-361
84. Le T, Wright EJ, Smith DM, et al. Enhanced CD4+ T-cell recovery with earlier HIV-1 antiretroviral therapy. N Engl J Med. 2013;368(3):218-230

85. Lederman MM, Funderburg NT, Sekaly RP, Klatt NR, Hunt PW. Residual Immune Dysregulation Syndrome in Treated HIV Infection. Adv Immunol. 2013;119:51-83
86. INSIGHT START Study Group. Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. N Engl J Med. 2015;373(9):795-807
87. Borges AH, Neuhaus J, Babiker AG, et al. Immediate Antiretroviral Therapy Reduces Risk of Infection-Related Cancer During Early HIV Infection. Clin Infect Dis. 2016;63(12):1668-1676
88. Grinsztejn B, Hosseinipour MC, Ribaudo HJ, et al. Effects of early versus delayed initiation of antiretroviral treatment on clinical outcomes of HIV-1 infection: results from the phase 3 HPTN 052 randomised controlled trial. Lancet Infect Dis. 2014;14(4):281-290
89. Yanik EL, Napravnik S, Cole SR, et al. Incidence and timing of cancer in HIV-infected individuals following initiation of combination antiretroviral therapy. Clin Infect Dis. 2013;57(5):756-764
90. Yanik EL, Napravnik S, Cole SR, et al. Relationship of immunologic response to antiretroviral therapy with non-AIDS defining cancer incidence. AIDS. 2014;28(7):979-987
91. Grulich AE, van Leeuwen MT, Falster MO, Vajdic CM. Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis. Lancet. 2007;370(9581):59-67
92. Brodin P, Jojic V, Gao T, et al. Variation in the human immune system is largely driven by non-heritable influences. Cell. 2015;160(1-2):37-47
93. Liston A, Carr EJ, Linterman MA. Shaping Variation in the Human Immune System. Trends Immunol. 2016;37(10):637-646
94. Murphy K. (2012) Janeway's Immunobiology, 8th Edition. New York, New York: Garland Science
95. Pothlichet J, Quintana-Murci L. The genetics of innate immunity sensors and human disease. Int Rev Immunol. 2013;32(2):157-208
96. Frank SA. (2002) Immunology and Evolution of Infectious Disease; Chapter 8: Genetic Variability of Hosts, . Princeton, New Jersey: Princeton University Press
97. Lane D, Levine A. p53 Research: the past thirty years and the next thirty years. Cold Spring Harb Perspect Bio. 2010;2(12):a000893
98. Lane DP. Cancer. p53, guardian of the genome. Nature. 1992;358(6381):15-16
99. Munoz-Fontela C, Macip S, Martinez-Sobrido L, et al. Transcriptional role of p53 in interferon-mediated antiviral immunity. J Exp Med. 2008;205(8):1929-1938
100. Munoz-Fontela C, Pazoss M, Delgado I, et al. p53 serves as a host antiviral factor that enhances innate and adaptive immune responses to influenza A virus. J Immunol. 2011;187(12):6428-6436
101. Rivas C, Aaronson SA, Munoz-Fontela C. Dual Role of p53 in Innate Antiviral Immunity. Viruses. 2010;2(1):298-313
102. Bouaoun L, Sonkin D, Ardin M, et al. TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data. Hum Mutat. 2016;37(9):865-876
103. Li FP, Fraumeni JF Jr, Mulvihil JJ, et al. A cancer family syndrome in twenty-four kindreds. Cancer Res. 1988;48(18):5358-5362

104. Lane DP, Crawford LV. T antigen is bound to a host protein in SV40-transformed cells. Nature. 1979;278(5701):261-263
105. Linzer DI, Levine AJ. Characterization of a 54K Dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. Cell. 1979;17(1):43-52
106. Eliyahu D, Raz A, Gruss P, Givol D, Oren M. Participation of p53 cellular tumour antigen in transformation of normal embryonic cells. Nature. 1984;312(5995):646-649
107. Parada LF, Land H, Weinberg RA, Wolf D, Rotter V. Cooperation between gene encoding p53 tumour antigen and ras in cellular transformation. Nature. 1984;312(5995):649-651
108. Eliyahu D, Michalovitz D, Oren M. Overproduction of p53 antigen makes established cells highly tumorigenic. Nature. 1985;316(6024):158-160
109. Finlay CA, Hinds PW, Levine AJ. The p53 proto-oncogene can act as a suppressor of transformation. Cell. 1989;57(7):1083-1093
110. Baker SJ, Fearon ER, Nigro JM, et al. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. Science. 1989;244(4901):217-221
111. Sarnow P, Ho YS, Williams J, Levine AJ. Adenovirus E1b-58kd tumor antigen and SV40 large tumor antigen are physically associated with the same 54 kd cellular protein in transformed cells. Cell. 1982;28(2):387-394
112. Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell. 1990;63(6):1129-1136
113. Werness BA, Levine AJ, Howley PM. Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science. 1990;248(4951):76-79
114. Friborg J Jr, Kong W, Hottiger MO, Nabel GJ. p53 inhibition by the LANA protein of KSHV protects against cell death. Nature. 1999;402(6764):889-894
115. Sato Y, Tsurumi T. Genome guardian p53 and viral infections. Rev Med Virol. 2013;23(4):213-220
116. Aloni-Grinstein R, Charni-Natan M, Solomon H, Rotter V. p53 and the Viral Connection: Back into the Future. Cancers (Basel). 2018;10(6):E178
117. Isobe M, Emanuel BS, Givol D, Oren M, Croce CM. Localization of gene for human p53 tumour antigen to band 17p13. Nature. 1986;320(6057):84-85
118. Kern SE, Kinzler KW, Bruskin A, et al. Identification of p53 as a sequence-specific DNA-binding protein. Science. 1991;252(5013):1708-1711
119. Khoury MP, Bourdon JC. The isoforms of the p53 protein. Cold Spring Harb Perspect Biol. 2010;2(3):a000927
120. Surget S, Khoury MP, Bourdon JC. Uncovering the role of p53 splice variants in human malignancy: a clinical perspective. Onco Targets Ther. 2013;7:57-68
121. Joerger AC, Fersht AR. The tumor suppressor p53: from structures to drug discovery. Cold Spring Harb Perspect Biol. 2010;2(6):a000919
122. Chen J, The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression. Cold Spring Harb Perspect Med. 2016;6(3):a026104
123. Rufini A, Tucci P, Celardo I, Melino G. Senescence and aging: the critical roles of p53. Oncogene. 2013;32(43):5129-5143

124. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol. 2010;2(1):a001008
125. Walker KK, Levine AJ. Identification of a novel p53 functional domain that is necessary for efficient growth suppression. Proc Natl Acad Sci USA. 1996;93(26):15335-15340
126. Sakamuro D, Sabbatini P, White E, Prendergast GC. The polyproline region of p53 is required to activate apoptosis but not growth arrest. Oncogene. 1997;15(8):887-898
127. Thomas M, Kalita A, Labrecque S, Pim D, Banks L, Matlashewski G. Two polymorphic variants of wild-type p53 differ biochemically and biologically. Mol Cell Biol. 1999;19(2):1092-1100
128. Dumont P, Leu JI, Della Pietra AC 3rd, George DL, Murphy M. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nat Genet. 2003;33(3):357-365
129. Beckman G, Birgander R, Sjalander A, et al. Is p53 polymorphism maintained by natural selection? Hum Hered. 1994;44(5):266-270
130. Sjalander A, Birgander R, Kivela A, Beckman G. p53 polymorphisms and haplotypes in different ethnic groups. Hum Hered. 1995;45(3):144-149
131. Sjalander A, Birgander R, Saha N, Beckman L, Beckman G. p53 polymorphisms and haplotypes show distinct differences between major ethnic groups. Hum Hered. 1996;46(1):41-48
132. Shi H, Tan SJ, Zhong H, et al. Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. Am J Hum Genet. 2009;84(4):534-541
133. Kay C, Jeyendran RS, Coulam CB. p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. Reprod Biomed Online. 2006;13(4):492-496
134. Chen H, Yang X, Wang Z. Association between p53 Arg72Pro polymorphism and recurrent pregnancy loss: an updated systematic review and meta-analysis. Reprod Biomed Online. 2015;31(2):149-153
135. Sinilnikova OM, Antoniou AC, Simard J, et al. The TP53 Arg72Pro and MDM2 309G>T polymorphisms are not associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. Br J Cancer. 2009;101(8):1456-1460
136. Baynes C, Healey CS, Pooley KA, et al. Common variants in the ATM, BRCA1, BRCA2, CHEK2 and TP53 cancer susceptibility genes are unlikely to increase breast cancer risk. Breast Cancer Res. 2007;9(2):R27
137. Sprague BL, Trentham-Dietz A, Garcia-Closes M, et al. Genetic variation in TP53 and risk of breast cancer in a population-based case control study. Carcinogenesis. 2007;28(8):1680-1686
138. Gaudet MM, Gammon MD, Bensen JT, et al. Genetic variation of TP53, polycyclic aromatic hydrocarbon-related exposures, and breast cancer risk among women on Long Island, New York. Breast Cancer Res Treat. 2008;108(1):93-99
139. Zhang X, Miao X, Guo Y, et al. Genetic polymorphisms in cell cycle regulatory genes MDM2 and TP53 are associated with susceptibility to lung cancer. Hum Mutat. 2006;27(1):110-117

140. Ren YW, Yin ZH, Wan Y, et al. P53 Arg72Pro and MDM2 SNP309 polymorphisms cooperate to increase lung adenocarcinoma risk in Chinese female non-smokers: a case control study. Asian Pac J Cancer Prev. 2013;14(9):5415-5420
141. Li Y, Chang SC, Niu R, et al. TP53 genetic polymorphisms, interactions with lifestyle factors and lung cancer risk: a case control study in a Chinese population. BMC Cancer. 2013;13:607
142. Mechanic LE, Bowman ED, Welsh JA, et al. Common genetic variation in TP53 is associated with lung cancer risk and prognosis in African Americans and somatic mutations in lung tumors. Cancer Epidemiol Biomarkers Prev. 2007;16(2):214-222
143. Hung RJ, Christiani DC, Risch A, et al. International Lung Cancer Consortium: pooled analysis of sequence variants in DNA repair and cell cycle pathways. Cancer Epidemiol Biomarkers Prev. 2008;17(11):3081-3089
144. Zhang L, Wang Y, Qin Z, et al. TP53 codon 72 Polymorphism and bladder cancer risk: a meta-analysis and emphasis on the role of tumor or smoking status. J Cancer. 2018;9(19):3522-3531
145. Gu Y, Zhou X, Zhang SL. Meta-analysis of an association of codon 72 polymorphisms of the p53 gene with increased endometrial cancer risk. Genet Mol Res. 2011;10(4):3609-3619
146. Liu KJ, Qi HZ, Yao HL, et al. An updated meta-analysis of the p53 codon 72 polymorphism and gastric cancer risk. Mol Biol Rep. 2012;39(8):8265-8275
147. Dahabreh IJ, Linardou H, Bouzika P, Varvarigou V, Murray S. TP53 Arg72Pro polymorphism and colorectal cancer risk: a systematic review and meta-analysis. Cancer Epidemiol Biomarkers Prev. 2010;19(7):1840-1847
148. Storey A, Thomas M, Kalita A, et al. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. Nature. 1998;393(6682):229-234
149. Hou J, Gu Y, Hou W, et al. P53 codon 72 polymorphism, human papillomavirus infection, and their interaction to oral carcinoma susceptibility. BMC Genet. 2015;16:72
150. Stoehr R, Weisser R, Wendler O, et al. P53 Codon 72 Polymorphism and Risk for Squamous Cell Carcinoma of the Penis: A Caucasian Case-Control Study. J Cancer. 2018;9(22):4234-4241
151. Zhou X, Gu Y, Zhang SL. Association between p53 codon 72 polymorphism and cervical cancer risk among Asians: a HuGE review and meta-analysis. Asian Pac J Cancer Prev. 2015;13(10):4909-4914
152. Li B, Wang X, Chen H, Shang LX, Wu N. TP53 codon 72 polymorphism and susceptibility to cervical cancer in the Chinese population: an update meta-analysis. Int J Clin Exp Med. 2015;8(6):9055-9062
153. Klug SJ, Ressing M, Koenig J, et al. TP53 codon 72 polymorphism and cervical cancer: a pooled analysis of individual data from 49 studies. Lancet Oncol. 2009;10(8):772-784
154. Dahabreh IJ, Schmid CH, Lau J, Varvarigou V, Murray S, Trikalinos TA. Genotype misclassification in genetic association studies of the rs1042522 TP53 (Arg72Pro) polymorphism: a systematic review of studies of breast, lung, colorectal, ovarian, and endometrial cancer. Am J Epidemiol. 2013;177(12):1317-1325

155. Menendez D, Inga A, Resnick MA. The expanding universe of p53 targets. Nat Rev Cancer. 2009;9(10):724-737
156. Wei CL, Wu Q, Vega VB, et al. A global map of p53 transcription-factor binding sites in the human genome. Cell. 2006;124(1):207-219
157. Smeenk L, van Heeringen SJ, Koeppel M, et al. Characterization of genome-wide p53-binding sites upon stress response. Nucleic Acids Res. 2008;36(11):3639-3654
158. Menendez D, Krysiak O, Inga A, Krysiak B, Resnick MA, Schonfelder G. A SNP in the fit-1 promoter integrates the VEGF system into the p53 transcriptional network. Proc Natl Acad Sci USA. 2006;103(5):1406-1411
159. Jordan JJ, Menendez D, Inga A, Noureddine M, Bell DA, Resnick MA. Noncanonical DNA motifs as transactivation targets by wild type and mutant p53. PLoS Genet. 2008;4(6):e1000104
160. Menendez D, Inga A, Resnick MA. The biological impact of the human master regulator p53 can be altered by mutations that change the spectrum and expression of its target genes. Mol Cell Biol. 2006;26(6):2297-2308
161. Tomso DJ, Inga A, Menendez D, et al. Functionally distinct polymorphic sequences in the human genome that are targets for p53 transactivation. Proc Natl Acad Sci USA. 2005;102(18):6431-6436
162. Menendez D, Shatz M, Azzam K, Garantziotis S, Fessler MB, Resnick MA. The Toll-like receptor gene family is integrated into human DNA damage and p53 networks. PLoS Genet. 2011;7(3):e1001360
163. Kaufmann SHE, Medzhitov R, Gordon S. (2004) The Innate Immune Response to Infection; Chapter 13: Toll-Like Receptors: Ligands and Signaling. Washington, DC: ASM Press
164. Bauer S, Muller T, Hamm S. Pattern recognition by Toll-like receptors. Adv Exp Med Biol. 2009;653:15-34
165. Garantziotis S, Hollingsworth JW, Zaas AK, Schwartz DA. The effect of toll-like receptors and toll-like receptor genetics in human disease. Annu Rev Med. 2008;59:343-359
166. Mills KH. TLR-dependent T cell activation in autoimmunity. Nat Rev Immunol. 2011;11(12):807-822
167. Armstrong L, Medford AR, Hunter KJ, Uppington KM, Millar AB. Differential expression of Toll-like receptor (TLR)-2 and TLR-4 on monocytes in human sepsis. Clin Exp Immunol. 2004;136(2):312-319
168. Taura M, Eguma A, Suico MA, et al. p53 regulates Toll-like receptor 3 expression and function in human epithelia cell lines. Mol Cell Biol. 2008;6557-6567
169. Shatz M, Menendez D, Resnick MA. The human TLR innate immune gene family is differentially influenced by DNA stress and p53 status in cancer cells. Cancer Res. 2012;72(16):3948-3957
170. Shatz M, Shats I, Menendez D, Resnick MA. p53 amplifies Toll-like receptor 5 response in human primary and cancer cells through interaction with multiple signal transduction pathways. Oncotarget. 2015;6(19):16963-16980
171. Heil F, Hemmi H, Hochrein H, et al. Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. Science. 2004;303(5663):1526-1529

172. Davila S, Hibberd ML, Hari Dass R, et al. Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis. PLoS Genet. 2008;4(10):e1000218
173. Oh DY, Taube S, Hamouda O, et al. A functional toll-like receptor 8 variant is associated with HIV disease restriction. J Infect Dis. 2008;198(5):701-709
174. Beima-Sofie KM, Bigham AW, Lingappa JR, et al. Toll-like receptor variants are associated with infant HIV-1 acquisition and peak plasma HIV-1 RNA level. AIDS. 2013;27(15):2431-2439
175. Mackelprang RD, Bigham AW, Celum C, et al. Toll-like receptor polymorphism associations with HIV-1 outcomes among sub-Saharan Africans. J Infect Dis. 2014;209(10):1623-1627
176. Dalgic N, Tekin D, Kayaalti Z, Cakir E, Soylemezoglu T, Sancar M. Relationship between toll-like receptor 8 gene polymorphisms and pediatric pulmonary tuberculosis. Dis Markers. 2011;31(1):33-38
177. Salie M, Daya M, Lucas LA, et al. Association of toll-like receptors with susceptibility to tuberculosis suggests sex-specific effects of TLR8 polymorphisms. Infect Genet Evol. 2015;34:221-229
178. Wang MG, Zhang MM, Wang Y, Wu SQ, Zhang M, He JQ. Association of TLR8 and TLR9 polymorphisms with tuberculosis in a Chinese Han population: a case-control study. BMC Infect Dis. 2018;18(1):561
179. Reference SNP (rs) Report – rs3764880 Available at: https://www.ncbi.nlm.nih.gov/snp/rs3764880
180. Reference SNP (rs) Report – rs3761624. Available at: https://www.ncbi.nlm.nih.gov/snp/rs3761624
181. Frank AK, Leu JI, Zhou Y, et al. The codon 72 polymorphism of p53 regulates interaction with NF-kB and transactivation of genes involved in immunity and inflammation. Mol Cell Biol. 2011;31(6):1201-1213
182. Detels R, Jacobson L, Margolick J, et al. The multicenter AIDS Cohort Study, 1983 to …. Public Health. 2012;126(3):196-198
183. Multicenter AIDS Cohort Study (MACS) Public Data Set. Available at: https://www.niaid.nih.gov/research/multicenter-aids-cohort-study-public-data-set
184. Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo CR Jr. The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. Am J Epidemiol. 1987;126(2):310-318
185. Seaberg EC, Wiley D, Martinez-Maza O, et al. Cancer incidence in the multicenter AIDS Cohort Study before and during the HAART era: 1984 to 2007. Cancer. 2010;116(23):5507-5516
186. Surveillance, Epidemiology and End Results (SEER) Program. Site Recode ICD-O-3 (1/27/2003) Definition. Available at: http://www.seer.cancer.gov/siterecord/icdo3_d01272003/
187. Aoshiba K, Tamaoki J, Nagai A. Acute cigarette smoke exposure induces apoptosis of alveolar macrophages. Am J Physiol Lung Cell Mol Physiol. 2001;281(6):L1392-L1401

188. Carnevali S, Petruzzelli S, Longoni B, et al. Cigarette smoke extract induces oxidative stress and apoptosis in human lung fibroblasts. Am J Physiol Lung Cell Mol Physiol. 2003;284(6):L955-L963
189. Bacon MC, von Wyl V, Alden C, et al. The Women's Interagency HIV Study: an observational cohort brings clinical sciences to the bench. Clin Diagn Lab Immunol. 2005;12(9):1013-1019

Curriculum Vitae

Filip Pirsl

7846 Heatherton Lane

Potomac, Maryland 20854

301-675-4604

fpirsl@gmail.com

Biographical Information

Born May 20, 1990 in Rijeka, Croatia

Education

ScM, Epidemiology
- Johns Hopkins Bloomberg School of Public Health – Baltimore, Maryland; August 2019

BSc, Physiology and Neurobiology
- University of Maryland, College Park – College Park, Maryland; May 2012

Research Experience

Contractor, Research Assistant
Kelly Government Solutions – Rockville, Maryland
June 2018-Present

Research Assistant
Experimental Transplantation and Immunology Branch,
National Cancer Institute, National Institutes of Health – Bethesda, Maryland
June 2017-June 2018

Post-Baccalaureate Cancer Research Training Award (CRTA) Fellow
Experimental Transplantation and Immunology Branch,
National Cancer Institute, National Institutes of Health – Bethesda, Maryland
June 2014-June 2017

Teaching Experience

Teaching Assistant
Johns Hopkins Bloomberg School of Public Health – Baltimore, Maryland

Data Analysis Workshop I – 140.613
Summer Institute, June 2019

Data Analysis Workshop II – 140.614
Summer Institute, June 2019

Stata Programming – 340.600
Fourth Term, March-May 2019

Epidemiologic Methods 3 – 340.753
Third Term, January-March 2019

Clinical Trials-Procedures, Design, and Interpretation of Results – 340.861
January-March 2019

Tutor
Johns Hopkins Bloomberg School of Public Health – Baltimore, Maryland

Statistical Methods in Public Health course series
September 2018-May 2019